

Construct Validity of High-stakes Tests

A Comparison of Finnish and Norwegian Written English Examinations

Andreas Nebdal



Masteravhandling ved Institutt for lærerutdanning og
skoleutvikling, Engelsk fagdidaktikk

UNIVERSITETET I OSLO

Høst 2012

© Andreas Nebdal

År: Høst 2012

Tittel: Construct Validity of High-stakes Tests: A Comparison of Finnish and Norwegian
Written English Examinations

Forfatter: Andreas Nebdal

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

Abstract

The main objective of this study is to investigate and compare the assessment of high-stakes, Vg3 level English examinations from Norway and Finland. The focus is on the differences in the approach to writing assessment in the two countries, and on the construct validity of the assessment procedure.

For this purpose a sample of six examinations from 2009, 2010 and 2011 was selected. Three Norwegian examinations from the Vg3 course “English Literature and Culture” were compared to three English matriculation examinations from the last year of upper secondary school in Finland. A qualitative content analysis of these tests was conducted to examine if the test content was in line with the test construct. They were analyzed to discover what they really tested, and how this corresponded with what the respective English syllabi describe. The analysis focused on three areas that could potentially pose a threat to the overall construct validity of the tests: construct-underrepresentation, construct irrelevance and scoring reliability.

First, the study showed a major difference in how writing was tested in the Norwegian and Finnish examinations. The Norwegian examinations consider aspects of academic writing skills, such as argumentative discourse and source referencing, while the Finnish examinations only test writing as the ability to produce a coherent text. Second, regardless of the differences, the results of the analysis indicate that the construct validity is threatened in both examination formats. There was little evidence of construct irrelevance, and nothing critical with regard to scoring reliability. However, construct-underrepresentation was found to be a threat; that is, the tests fail to include important aspects of what the syllabi prescribe. In Norway, it was the topical issues of culture and society mentioned in the Norwegian syllabus that was found to be under-represented in the examinations. For the Finnish examinations, it was mainly a failure to test genre-specific writing, despite this being mentioned as a skill in the Finnish English syllabus.

The validity of the findings is restricted to the six examinations in the sample, and relies somewhat on the subjective judgment of the researcher, but the findings should open for discussion about the national examination systems in general. The study suggests that care be taken to consider the content of future examinations, ensuring that what they test is better in accordance with the respective syllabi in both countries.

Sammendrag

Målet med denne studien er å undersøke og sammenligne norske og finske avgangseksamener på Vg3-nivå i Engelsk. Fokus er på forskjellene i tilnærmingen til vurdering av skrijving i de to landene, og på vurderingsprosedyrens konstruktvaliditet.

Med det som hensikt ble det gjort et utvalg av seks eksamener fra 2009, 2010 og 2011. Tre norske eksamensoppgaver fra Vg3 kurset «Engelsk litteratur og kultur» ble sammenlignet med tre såkalte «studenteksamener» fra siste året i videregående skole i Finland. En kvalitativ innholdsanalyse av disse prøvene ble utført for å undersøke om prøveinnholdet stemte overens med prøvekonstruktet. De ble analysert for å finne ut hva de egentlig testet, og hvordan dette korresponderte med det som beskrives i de respektive læreplanene. Analysen fokuserte på tre områder som potensielt kunne utgjøre en trussel for vurderingsprosedyrens konstruktvaliditet: konstrukt-underrepresentasjon, konstruktirrelevans og reliabilitet.

Først og fremst viste studiet en betydelig forskjell på hvordan skrijving ble testet i de norske og de finske eksamenene. De norske eksamenene inkluderer aspekter av akademiske skriveferdigheter, slik som argumenterende diskurs og kildereferanser, mens de finske eksamenene kun tester skrijving som evnen til å produsere en sammenhengende tekst. For det andre, uavhengig av disse forskjellene, indikerte resultatet av analysen at konstruktvaliditeten er truet i begge eksamensformatene. Det var lite belegg for konstruktirrelevans, og intet kritisk å finne med tanke på reliabilitet. Imidlertid viste analysen konstruktunderrepresentasjon som en mulig trussel. Det vil si at prøvene unnlater å inkludere viktige aspekter av det læreplanen foreskriver. I Norge var det kultur- og samfunnsemnene nevnt i læreplanen som var underrepresentert i eksamensoppgavene. Når det gjald de finske eksamenene var det hovedsakelig mangel på testing av sjangerspesifikk skrijving til tross for at det nevnes som en ferdighet i den finske læreplanen.

Validiteten til funnene er begrenset til de seks eksamenene i utvalget, og beror til dels på subjektiv bedømmelse, men de skulle likevel kunne åpne for diskusjon rundt de nasjonale eksamenssystemene generelt. Studien foreslår at innholdet i fremtidige eksamener blir overveid for å sørge for at det som testes er i bedre overensstemmelse med de respektive læreplaner i begge land.

Acknowledgements

I would like to thank my wife, Jonna-Jasmina, for being an immense support for me, as always, so also during the time of working with this thesis. Especially during the last months before the final deadline, she has often had to be alone taking care of our son, and I am grateful beyond words. My supervisor, Associate Professor at the University of Oslo, Glenn Ole Hellekjær, also deserves acknowledgement. I would like to express my deepest appreciation for his comments, suggestions, help and motivation. Finally, I am also convinced that I would not be where I am now, if it were not for God's help through everything. He has given my life an ultimate purpose, which in turn gives everything in life a purpose.

Table of contents

1	Introduction	1
1.1	Background.....	1
1.2	Norway and Finland	2
1.3	Pilot study	3
1.4	Other related studies	4
1.5	Research statement	6
1.6	Brief overview of the thesis	8
1.7	Comments on terminology	8
2	Assessment.....	9
2.1	Defining assessment and the use of assessment	9
2.1.1	The two elements of assessment	9
2.1.2	The purpose and use of assessment.....	11
2.2	Types of assessment	12
2.2.1	Basic distinctions.....	12
2.2.2	Types of language tests	15
2.2.3	Language test design	16
2.3	Defining the construct	18
2.4	Defining construct validity	19
2.4.1	The simple and outdated view.....	19
2.4.2	The broad and current view.....	20
2.4.3	Validation	21
3	Writing	24
3.1	Types of writing.....	24
3.2	Cognitive model	27
3.2.1	Input and process.....	27
3.2.2	Knowledge telling and knowledge transformation	28
3.3	Communicative competence.....	29
3.3.1	Linguistic competence.....	30
3.3.2	Socio-linguistic competence	31
3.3.3	Pragmatic competence.....	31
3.4	Summing up.....	32

4	Norway vs. Finland	33
4.1	A comparison of the school and examination systems.....	33
4.1.1	National curricula.....	33
4.1.2	English in primary and lower secondary school	34
4.1.3	Courses in upper secondary school:	36
4.1.4	Examinations:.....	37
4.2	The English syllabi	38
4.2.1	The Norwegian English syllabus.....	38
4.2.2	The Finnish English syllabus	40
4.2.3	The construct definitions of writing	42
5	Method	44
5.1	Chapter overview.....	44
5.2	Research design	44
5.3	Sample	44
5.4	Content analysis.....	45
5.5	Validity and limitations	49
6	Results of the content analysis	52
6.1	Overview of chapter	52
6.2	Description of the examination formats	52
6.2.1	The Finnish English examinations	52
6.2.2	The Norwegian English examinations	54
6.2.3	Comparison of the two examination formats	55
6.3	Description of the assessment documents	55
6.3.1	The Finnish assessment documents.....	55
6.3.2	The Norwegian assessment documents.....	57
6.3.3	Comparison of assessment documents.....	59
6.4	Test requirements: What do the tests measure?.....	60
6.4.1	Using the cognitive model.....	60
6.4.2	Reading comprehension	61
6.4.3	Topical knowledge	64
6.4.4	Language knowledge.....	68
6.4.5	Cognitive level	71
6.4.6	Summing up so far: Comparison of requirements	73

6.5	The relationship between tests and construct: Do they measure what they should? .	75
6.5.1	Construct irrelevance.....	75
6.5.2	Construct under-representation	81
6.6	Scoring reliability: Do they measure reliably?	86
6.6.1	The Finnish examinations	86
6.6.2	The Norwegian examinations.....	87
6.6.3	Comparison of scoring reliability.....	88
7	Discussion	90
7.1	Chapter overview.....	90
7.2	Summary of the findings	90
7.2.1	The different approaches to writing assessment	90
7.2.2	The threats to construct validity	91
7.3	Validity of the findings.....	93
7.4	Discussing the different approaches to writing assessment.....	94
7.4.1	Testing academic writing	94
7.4.2	Testing communicative competence	96
7.4.3	Testing topical knowledge	97
7.5	Discussing construct validity.....	99
7.5.1	Brief review of construct validity and validation.....	99
7.5.2	Reliability	99
7.5.3	Construct mismatch.....	100
7.6	Test consequences	104
7.7	Discussion summary.....	105
8	Conclusion.....	107
8.1	Further research	107
8.2	Implications	109
	References	111
	Appendix 1	116
	Appendix 2	136
	Appendix 3	156
	Appendix 4	176
	Appendix 5	190
	Appendix 6	206

List of tables

Table 1: Types of assessment.....	15
Table 2: Language Task Characteristics	16
Table 3: Scoring criteria for written production in Finnish English examinations.....	56
Table 4: Scoring criteria for top grades (5-6) in the Norwegian English examinations	58
Table 5: Reading requirements in the Finnish English examinations.....	61
Table 6: Reading requirements in the Norwegian English examinations	63
Table 7: Topical knowledge required in the Finnish composition tasks.....	65
Table 8: Topical knowledge required in the Norwegian examinations	66
Table 9: Genre specifications in the Finnish English examinations	69
Table 10: Genre specifications in the Norwegian English examinations.....	70
Table 11: Types of writing tasks in the Finnish examinations.....	71
Table 12: Types of writing tasks in the Norwegian English examinations.....	72
Table 13: Comparison of what the Finnish and Norwegian English examinations measure...	73
Table 14: Construct irrelevance for the Finnish English examinations	76
Table 15: Construct irrelevance for the Norwegian English examinations	78
Table 16: Number of tasks where each part of the construct is measured in the Finnish English examinations.....	82
Table 17: Number of tasks where language and language learning aims are measured in the Norwegian English examinations.....	83
Table 18: Number of tasks where communication aims are measured in the Norwegian English examinations	84
Table 19: Number of tasks where culture, society and literature aims are measured in the Norwegian English examinations.....	85

List of figures

Figure 1: Validation process of a test.....	22
Figure 2: Distribution of language levels in Finnish compulsory education	35
Figure 3: Illustration of construct under-representation and construct irrelevance	47

1 Introduction

1.1 Background

Imagine that there is a large group of people who claim that they know English fairly well. Imagine that you need the five best of them for a special program in your profession. How would you find out which ones to choose? If you would use a language test, how would that test look like? How could you be sure that the test would give you the right results? What does it really mean to know a language, and how can it be fairly and properly tested? There seem to be many opinions on these matters, and even within the public school systems of the Nordic nations, there are great differences in the approach to language testing. Assessing language is useful for more than selecting the best five of a large group, but the issues remain the same: What should be assessed? What is a good way to do such an assessment? And how can it be done within the framework of the school system?

Typically, language assessment in school is done through a final examination after completing a language course. Because of family relations with Finland, I have learned of many interesting differences between the Norwegian and Finnish education systems. In fact, when I discovered the difference between the English examinations in the two countries, I was curious to know how this could be so. How could two so dissimilar test designs both be accepted ways to assess English?

For me, these types of questions motivated the topic of my paper. The general questions of how language skills are measured triggered a philosophical interest in me; a need to understand more of the nature of language and language use. Next, the questions of how the quite different tests in Norway and Finland have developed triggered a need to understand how these examination systems are constructed. I will not enter into the world of philosophy, or into educational politics in this paper, but these are factors that motivated my study, and that also put the thesis in a larger perspective. Although I am only looking at a sample of tests and discussing their differences, it is in my intention to also draw attention to the topics of how schools today best can assess foreign language ability. As a preliminary research statement, one could say that this is a comparative study of upper secondary school level written English examinations used in Norway and Finland.

1.2 Norway and Finland

Since this is a comparative study, I find it important to give some comments on the comparison from the start. My primary reason for choosing to compare Norway and Finland was personal, myself being Norwegian, and my wife Finnish. However, I am not the only one interested in studying the Finnish education system. During the last decade, Finland and its school system have been on the lips of many reporters and experts. In “The Programme for International Student Assessment” (PISA) from 2000 to 2006, Finland was ranked as number one in all disciplines: reading literacy, mathematics and science (Organisation for Economic Co-operation and Development, 2001, 2004, 2007). This has caused many to wonder how the Finns organize their schools, giving me all the more motivation for this thesis.

Foreign languages, however, are not a part of the PISA study (Store norske leksikon, n.d.). And, even though foreign languages are highly emphasized in the Finnish education system, it is possible that Finnish students would not excel as much in that area. My general impression from being in Finland is that their level of English is not quite on a par with the level in Norway. One may assume that this is because the Finnish language is much further from English than the Norwegian language is. Finnish belongs to the Finno-Ugric language family, which is not even a part of the Indo-European languages. By contrast, Norwegian and English are both Germanic languages, and therefore very closely related. This difference could arguably discredit a comparison of English assessment in these two nations, and it should also be kept in mind when reading the thesis. At the same time, Finnish is not the only official language in Finland. Six percent of the Finnish population have Swedish as their mother tongue (Giverholt, 2009), and the Finland-Swedes follow to the same education system as the rest and have the same examinations. In other words, the comparison is relevant since Swedish-speaking and Norwegian-speaking students practically have the same point of departure.

Finally, it should be mentioned that analyzing the Norwegian examinations only might have been interesting as well, but by seeing the Norwegian tests in light of the Finnish ones, and vice versa, I hope to create a contrastive perspective which will better highlight the particular features of each national examination system.

1.3 Pilot study

Before I started my master thesis, I attended a course called “Assessment and Testing in English as a Foreign Language” at the University of Oslo, and that was when I first realized the major challenges of language assessment. As an assignment for that course, I conducted a small study quite similar to this one. My sample was one Norwegian and one Finnish test from the autumn of 2010, and I analyzed them using the following checklist questions:

- 1) What method and skills do the tasks require?
- 2) Is the language ability construct¹ for this test clearly and unambiguously defined?
- 3) Is the language ability construct for the test relevant to the purpose of the test?
- 4) To what extent does the task and scoring procedures reflect the construct definition, and will they help us make the desired interpretations about test takers’ language ability?

With this I wanted to investigate the construct validity (for definition, see section 2.4) and find the reasons for the differences in the tests. First of all, I found that despite the different task types, the two tests measured many of the same aspects. The most prominent difference was that the Norwegian test required more knowledge about topical issues of culture, literature and society. For both examinations the syllabi provided clear test constructs that were also relevant for their purpose. However, I discovered a mismatch between the communicative view of language described in the Finnish syllabus and the effect of the multiple-choice tasks. In a multiple-choice task, it is only one correct form that is accepted. However, if communication is the primary goal, incorrect grammar does not always matter as much.

This led to a discussion about construct validity, as well as about the importance of a valid construct. In the real world, there are many factors to consider when developing a test, so I also discussed the notion of balancing different test qualities like validity and reliability (for definitions, see section 2.4). Finally, I briefly touched upon the issue of washback, that is, consequences of the tests on teaching and learning. In light of such consequences, I questioned the use of multiple-choice tasks with regard to how it influences English teaching in school. In my conclusion, I suggested that “perhaps my master thesis in one year could enhance this study”, which is exactly what I decided to do.

¹ For definition, see section 2.3

1.4 Other related studies

Studying the construct validity of language tests is a common phenomenon, and has been done for decades. The theories behind validity and validation, as well as the practical side of test development, have been studied and developed throughout the years. Thirty years ago, L2 assessment validation research was mostly concerned with test outcomes and the reliability and inter-correlation of subtests (Cohen, 2006). There was less focus on how the test takers arrived at their responses and how it related to the abilities that were to be tested (Cohen, 2006). Among current research on construct validity of language tests, there is still a considerable proportion doing large-scale statistical analyses of well-known high-stakes tests. Some compare two test formats to see if they measure the same (Zahedi & Shamsaee, 2012), while another variant is comparing different types of tasks to see if they measure the same (Guo, 2011). Lazaraton and Taylor (2007) point out the need for qualitative research methods like “Discourse and Conversation Analysis, Observation Checklists and Verbal Protocol Analysis” (pp. 113-114). These are more in line with my study, and have also grown to be common methods (O’Sullivan, Weir, & Saville, 2002; Cumming, Grant, Mulcahy-Ernt, & Powers, 2004; Plakans, 2009).

Most relevant for the present thesis are studies related to examinations used in Norway and Finland. Finding such studies, however, was not so easy, and others have noticed the same: “On the whole, there is not much systematic research into the Matriculation Examination, which is not unusual in the European context of national examinations” (Lindström, 1998 as cited in Huhta, Kalaja, & Pitkänen-Huhta, 2006, p. 329).

I was only able to find a few studies relevant to the Finnish English examinations. Huhta, Kalaja and Pitkänen-Huhta (2006) looked at different students' preparation to the English matriculation examination, using an oral diary of their thoughts and experiences. The purpose of the study was to gain insight into the meaning the students give the test. In a validity perspective, one may say that it deals to some degree with consequential validity, in particular social consequences of high-stakes tests (Huhta, Kalaja, & Pitkänen-Huhta, 2006). Two other studies have looked more directly at the consequential effects of the English matriculation examination on English textbooks, also called washback (Vainio 2009; Pietilä, Taanila-Hall, & Vainio, 2009). One of them (Vainio, 2009) found that the textbooks did not reflect the multiple-choice format of the test, while the other (Pietilä, Taanila-Hall, & Vainio, 2009)

found that the lack of oral exercises in the textbooks might be a washback effect of the matriculation examination.

The current Norwegian course English Literature and Culture (Utdanningsdirektoratet [Udir], 2006) is fairly recent, so it came as no surprise that no studies of those examinations were to be found. Instead, I will mention a couple of other studies related to the assessment of English in Norway. One is a recent master's level study of the English oral examination formats used at the Vg1 level (Yildiz, 2011), which found extensive variation from county to county. Contrary to the written English examinations, the oral ones are not administered nationally, which, at least in part, can account for this variation. While these findings may not be generalized to all the oral examinations at the Vg1 level, it still shows cause for concern about their construct validity.

Even closer to my field of study are two master's theses that analyzed academic writing in Norwegian upper secondary school (Sparboe, 2008; Shirazi, 2010). The studies included analyses of the curriculum, a selection of textbooks, a selection of final examinations and interviews with teachers. The earliest one (Sparboe, 2008) was written before the new curriculum in 2006, the Knowledge Promotion (LK06), and found that academic writing instruction was largely overlooked in Norwegian upper secondary school. The more recent thesis (Shirazi, 2010) conducted the same study within the framework of the new curriculum, and found that the situation had changed. LK06 includes academic writing features, which arguably better prepares the pupils for higher education. The final examinations have apparently followed suit and also include this aspect (Shirazi, 2010).

Even more recently, an article submitted to *Acta Didactica* by Sigrid Ørevik (2012) presents an analysis of genre specifications in Norwegian English examinations from 1996 to 2011. It examines the genre patterns in sixteen Vg1 level examinations. Ørevik found that genres asked for in the writing tasks have not changed much over the years, and that they are often characterized by unclear and vague genre instructions (Ørevik, 2012).

As in Finland, there is also a study in Norway that has investigated possible washback effects, i.e. consequences on teaching, of Norwegian English examinations (Ellingsund, 2009). At the Vg1 level, twelve teachers were interviewed to see if their classroom instruction was

influenced by the final examinations, and around half of them claimed that it was, to various degrees.

Lastly, although it is not so recent, I will present a study by Synnøve Pettersen (2001) which resembles mine in many ways. It is a study investigating the correspondence between English Vg1 examinations and syllabi in Norwegian upper secondary school from 1976 to 2000. The results are not valid for the present Norwegian curriculum, but she found an increasing focus on communicative competence, where the more recent examination tasks provided more natural contexts and authentic situations for writing. Moreover, the most recent examinations in the study were found to be more consistent with the syllabus targets, also including the targets concerning cultural knowledge (Pettersen, 2001).

To sum up, it seems that although validation studies in general are fairly numerous, studies on the Nordic English examinations are few and far between. In Finland, there has been some emphasis on test preparation, and how it affects the lives of the students, as well as washback effects, but nothing on the quality of the test content. For the Norwegian situation, there are quite recent studies concerned with oral testing, academic writing, washback and use of genres in upper secondary English examinations. There is also a less recent study comparing syllabus targets and what the English examinations measure, which is in some ways similar to this thesis.

To my knowledge, there is no other recent study, Norwegian or Finnish, that considers the English examinations with regard to construct validity as I do in the present study. Neither have I found any comparative studies across the Nordic nations with this focus. It is therefore my aim to fill in that gap, and provide some new information about the construct validity of English writing assessment in Norway and Finland, hence my research statement in the next section.

1.5 Research statement

The research statement of this study is:

How does the construct validity of writing assessment compare between Finnish and Norwegian English examinations at the highest level of upper secondary school?

Put more simply, the question is how Finnish and Norwegian Vg3 level² English tests assess writing ability. In more technical terms, it involves a *comparison of writing assessment* in terms of *construct validity*. The concept of construct validity will be explained in detail in chapter 2, and is the most important quality of a test. It has to do with whether a test tests what it is intended to measure, so that the test scores can be used for what they are intended to be used for. In fact, a test might not test what we think it does, and, as a consequence, the scores may be used for invalid purposes. When testing writing ability, this might happen because writing is such a comprehensive term, as chapter 3 will also explain. The ones responsible for the examinations in upper secondary school in Norway and Finland need to define writing in their English syllabi. But the syllabi might define it in one way, and the tests test it in another. This was what I wanted to investigate, and in my investigation of the tests, three core questions have guided the analysis:

- What do the tests measure?
- Do they measure what they should?
- Do they measure reliably?

Understanding what the tests measure required an analysis of each task in the tests, and enabled me to compare the different approaches of the Norwegian and Finnish examinations. It also lays the foundation for the next question, of whether the examinations measure what they should – the abilities described in the respective English syllabi. In order to know whether they do so or not, it was necessary to first find what it is they actually measure. Next, I had to examine the syllabi and see whether the tests corresponded with these. Construct validity is dependent on a match between the two.

Construct validity is also dependent on reliable scoring procedures, hence the last question. If the test scores depend on other factors than the abilities of the test takers, it also weakens the validity of the test. The analysis was therefore supplied with information about the scoring procedures and rater guidelines of the examinations in question. The method is more fully described in chapter 5, which also contains the full list of checklist questions used for the analysis. The next section provides an outline of all the chapters in the thesis.

² Vg3 is the highest level of upper secondary school. See section 4.1 for comparison of school systems

1.6 Brief overview of the thesis

The two following chapters, chapter 2 and 3, will together provide the theoretical background for this study. Chapter 2 deals with types of assessment in education, and particularly assessment of foreign languages. It also brings in the challenges involved in measuring language ability, and defines the measurement terms *construct* and *construct validity*. Chapter 3 looks into the field of writing and language theory. A cognitive model of writing is outlined, as well as the language view known as communicative competence.

All through the thesis I will keep a comparative perspective, where the similarities and differences between Norway and Finland will be brought to attention. Chapter 4, in particular, is set out to outline and compare the two national educational systems, and particularly the English syllabi for the highest level of upper secondary school. At the end of the chapter, construct definitions of the national examinations are also presented.

Chapter 5 explains the research design and the size and selection of the sample. The procedure of my analysis is included, as well as comments on the validity of the study. This is followed by the results of the analysis in chapter 6, which are summarized and discussed in chapter 7. Finally, the conclusion is found in chapter 8 and includes some of the implications of the findings and suggestions for further research.

1.7 Comments on terminology

The most essential terms for this study will be defined in the following two chapters. There are only a few comments necessary regarding use of other terminology. In this paper, I have seen no need to differentiate between *test* and *examination*. Instead I have used them both about the examinations in the sample for the sake of variation. Closely related to this is the word for the person who takes the test. Since that is referred to often, I have varied between the following expressions: *test taker*, *testee*, *examinee* and *candidate*, without any difference in meaning. These are all terms I have come over in the assessment literature I have been studying. As a final comment, I may also mention that I have mostly used the word *student* for learners in upper secondary school. Other documents, however, might sometimes also refer to them as *pupils*, in which case I have kept the original term in my quotes.

2 Assessment

As the first of two theory chapters, this one will deal with the area of assessment, and some of the basic terminology and concepts related to this field. It will start defining educational assessment and separate it from what is defined as assessment use. Secondly, based on this distinction, the basic types of assessment will be described. The focus will further be put on formal assessment of language, that is to say, the part of assessment known as testing and the different types of language tests. Next, it will give an overview of types of test tasks, and outline the options for test design. Finally, there will be two sections on two important assessment concepts. The first of them concerns the construct of a test, and how it is defined. The second is an introduction to the current view of construct validity.

2.1 Defining assessment and the use of assessment

2.1.1 The two elements of assessment

One will often associate the word assessment with tests in school. A test is certainly a type of assessment, but assessment is much broader. In its general sense, it is not even something only reserved for educational settings, as long as the two defining elements are included, which will be explained below. However, the focus of this paper will be on assessment for educational purposes.

If a teacher gives his class an exercise and notices that many of the students seem confused, he will probably conclude that the exercise is hard to understand. Although this is not what we call a test, it is also a type of assessment. Different types of assessment will be discussed in the following section, but there are two elements involved in all types. In this example, when the teacher assessed the situation, he did two things. First, he observed the class and noticed the confusion. Second, he interpreted the observation and made a conclusion based on it. Observation and interpretation are the two key elements of assessment, but since it is often impossible to observe directly what we are interested in, a more precise terminology is needed. Harlen (1994), as cited in Broadfoot (2007, p.4), uses the following definition:

“assessment is the process of firstly gathering evidence, and secondly interpreting that evidence in the light of some defined criterion in order to form a judgment”. The first element here is not called observation but “gathering evidence”. Bachman uses yet another expression in his definition: “Assessment is the process of *collecting information* [emphasis added] about something that we're interested in” (Bachman, 2004b, as cited in Bachman & Palmer, 2010, p.20).

In education, we are usually interested in the students' knowledge and abilities, which are not directly observable. This is why we need to gather evidence of such knowledge, or collect some type of information which says something about these abilities. For instance, a student may know a lot about World War I, but there is no way to look into his brain and observe this knowledge. He needs to be given a way to show evidence of such knowledge, and this must be collected and interpreted by an assessor. This can be done by observing classroom activities, by giving a test, or by other means. Regardless of the means used, it will never provide a complete and accurate picture of the knowledge the student possesses. The information must still be considered evidence of such knowledge, and good assessment will provide sufficient evidence. In our first example, the teacher was interested in whether the given exercise was too difficult or not. The confused face expressions were considered evidence that the exercise was difficult, and this was therefore his conclusion. If the teacher decided to simplify the exercise as a result of this conclusion, it would be use of the assessment. More details on the use of assessment, however, will be covered below.

To sum up, assessment is defined as the exercising of the two elements mentioned:

(1) Gathering evidence and (2) interpretation of that evidence. Gathering evidence is the first step, where some kind of performance or behavior on part of the assessee will be recorded or observed. Bachman and Palmer's (2010) term for this step helps to clarify what this means in the case of a test. They use the term “test taker's performance”, but they specify that the term “test taker” is only used for convenience since assessment does not always include a test.

However, in the case of a test, it is the test taker's performance on that test which is the evidence. They also use the term “assessment task” about the means of gathering evidence.

This assessment task is used in order to elicit the desired performance by the test taker, or one

might say in order to gather evidence. Such an assessment task may be a test, but there are also other means of gathering the evidence which will be a topic in the next section.

Interpretation is the second step of assessment. Interpretation is making meaning of the data gathered in step one. In a testing situation, this will usually happen through some kind of scoring procedure and result in a description or a grade or both. Otherwise, it may happen momentarily and less systematically as in the example of the confused students. Having established the definition of assessment as involving these two steps, we may look at the purpose and use.

2.1.2 The purpose and use of assessment

Prior to an assessment, there may be a planning phase where the means of assessment are produced. After an assessment, there is another phase where the outcome of the assessment is being used for something. If the information provided by the assessment were not to be used for anything, there would be no need for an assessment at all. Decisions made on the basis of the assessment are the purpose and use of assessment. Therefore, if the teacher would decide to use another exercise based on his interpretation of his confused-looking students, this would not be a part of the assessment itself. This would be a use of his assessment, and the very purpose why the assessment took place. If there is a need to make a decision based on the language abilities of a group of people, an assessment providing such information must be conducted. To further illustrate the relation between assessment and assessment use, it is useful to look at how Bachman and Palmer (2010) describe it. They list up five different elements in a series, where each element in the series is derived from the preceding element:

- (1) Test taker's performance
- (2) Assessment record
- (3) Interpretation about test taker's ability
- (4) Decisions
- (5) Consequences. (p. 23)

The first step has already been mentioned. It concerns the first element in the assessment process, the gathering of evidence through some performance by the learner. Steps two and three are the scoring and interpretation, the second element in the assessment process. One may notice that this description includes assessment record before the actual interpretation of the evidence, but these two points may be considered one for our purpose. Thus step one to three are part of our definition of assessment. Step four, however, decisions, is the use of

assessment. Just as the interpretation is based on the performance, the decisions are based on the interpretation. This may for instance be decisions about admission to an educational course, or smaller decisions involving how to improve the teacher's instruction in class. Last, step five brings in the aspect of consequences of these decisions. That is a highly discussed area in assessment, but will only be briefly covered here. The main point of this first section has been to clearly define these separate steps because they are essential to an understanding of the concept of construct validity.

2.2 Types of assessment

2.2.1 Basic distinctions

As we have seen, assessment includes a variation of activities that all have in common the two basic elements mentioned above. Both formal examinations and sometimes classroom activities may be called assessment, but they are not the same type of assessment. It is possible to divide assessment into types according to the way in which each of the elements in the process vary. A closer look at each of these four steps, and some common distinctions and variations within each one, will provide a relevant overview of the main types of educational assessment.

Gathering evidence

The most basic distinction in types of assessment is based on whether the evidence is gathered formally or informally. Another expression for this, used by Bachman and Palmer (2010), is whether the assessment is done explicitly or implicitly. Thus we divide between assessment where the learners are largely unaware of the assessment, and assessment where both the learners and the teacher are aware of it. The former is informal or implicit, while the latter is called formal or explicit. In the case of implicit mode, the assessment is not so clearly distinct from the process of teaching itself. It may happen during a class of normal teaching instructions, and even be part of the instructions themselves. Bachman and Palmer (2010) specify that this type of assessment takes place in a continuous, instantaneous and cyclical manner. This may for instance be in a class if a student asks the teacher a question. The question may cause the teacher to decide to change or revise his lesson plan. This change may again cause new questions to which the teacher responds and which may influence the

following course of the lesson. It is a continuous process where the teacher constantly receives information from classroom dialog and observation (assessment), and decides to adjust his teaching based on this information (assessment use). Even though this is very informal, and may even happen while the teacher is unaware that assessment is taking place (Bachman & Palmer, 2010), it still fits our definitions of assessment and assessment use.

If, on the other hand, the students are made aware of an assessment situation, it is explicit or formal assessment. Gathering evidence in an explicit or formal way is what we call testing. Tests are means of gathering evidence while everyone is fully aware of the test situation. In this case, it is to a greater extent separated from the teaching. So although there may be numerous ways of gathering evidence, one basic distinction is between using formal tests and using other means. A closer description of different types of tests will be presented below, with focus on language tests.

Interpretation

The main distinction of gathering evidence was explicit or implicit. If it happens implicitly, the process is instantaneous, and the assessment and the decisions may be done in a moment (Bachman & Palmer, 2010). In such cases, it may be hard to categorize types of interpretation since it is very individual, and often not a thorough, systematic procedure. However, when interpreting explicitly collected evidence, like test results, it is possible to categorize different types of procedures. One common distinction is whether to do the scoring based on a normal distribution or based on a set of criteria. Using normal distribution is not uncommon for large standardized tests, and means that the test takers are competing with one another. The scores of all the test takers will then be gathered and compared. It is not decided beforehand what score will correspond to what grade because it depends on the result. Instead the percentage of how many will achieve each of the grades is given by the statistical normal distribution curve. In comparison, if the interpretation is based on criteria, each test taker will be measured according to these, and assessment will not be affected by other test takers' performance. This means that the test developers have settled on set criteria in advance for how to achieve the different grades. In contrast to norm-based tests, this means that everyone may theoretically achieve the highest grade as long as they all meet the set criteria.

Decisions

Decisions in language assessment are made about individuals and programs, and can be roughly divided into summative and formative decisions (Bachman & Palmer, 2010).

Formative decisions are decisions to make modifications in the teaching or learning process. The teacher may decide to change his instructions, or the student may be given feedback in order to change his learning strategy. Summative decisions are about the final outcome of a course, typically about selection or certification (Bachman & Palmer, 2010). In order to make this type of decisions, the assessment mode must be explicit. In implicit assessment, one may only make formative decisions, but in explicit assessment, both are possible. There are many parallels here to the distinction between “assessment of learning” and “assessment for learning”. The purpose of summative decisions is to provide information for stakeholders; it is assessment *of* the learning that has taken place, and informing some stakeholder of the result. The purpose of formative decisions is to improve the learning process, assessment *for* learning. It means that we assess rather to promote learning than to check or inspect what is learned.

Traditionally, assessment research has mostly been concerned with the qualities of large-scale examinations, assessment of learning. But it is also worth noticing the recent trends of learning-related assessment:

As mentioned at the beginning of the article, the biggest changes in language assessment in recent decades have been about the scope of the field of language testing. Earlier, it was considered to be mostly about large-scale tests with strong emphasis on statistical analysis. While that is currently one of the main areas of activity and research, another equally important strand is learning-related assessment. The formats of assessment that are relevant here include various kinds of self-evaluation and peer evaluation, portfolio assessment, learning diaries, etc. In terms of testing theory, the development has meant that language testers have had to rethink their assumptions about assessment as an activity, as well as the quality criteria that apply in different contexts (Luoma, 2002, para. 19).

Learning-related assessment, or assessment for learning, has become such an important field that it has consequences for all of assessment theory. It questions the very nature of assessment, and the role it has in society. For this reason, it is worth keeping in mind although this paper is concerned with large-scale summative assessments.

Consequences

When carrying out an assessment, and making decisions based on it, there will always be consequences (Bachman & Palmer, 2010). Some of these consequences may be those intended by the assessment user, but there may also be unintended consequences. Details of the consequential aspects of assessment are beyond the scope of this thesis. It is usual to divide formal assessment into high-stakes and low-stakes assessment (Simensen, 2007). This is a question of whether the consequences of the decisions are of major or minor significance in a greater social perspective. Implicit assessment will always be relatively low-stakes, where small decisions in the classroom may have minor consequences for a class or an individual for a short time. Formal national examinations, however, are usually high-stakes, where the outcome may have consequences for a vast number of people, for instance in terms of access to higher education.

In conclusion, table 1 below shows how these four steps of assessment and assessment use make it easy to categorize the different types of assessment. The next section will take a closer look at the left column, on explicit assessment, or tests, in the field of language testing.

Table 1: Types of assessment

1. Gather evidence	Explicit Formal Tests	Implicit Informal Non-tests
2. Interpret evidence	Norm-based / Criterion-based	Individual
3. Make decisions	Summative / Formative	Formative
4. Consequences	High-stakes / Low-stakes	Low-stakes

2.2.2 Types of language tests

Categorizing different types of language tests is usually based on the purpose of the test. Simensen (2007) lists six common purposes of language tests, and different test types that are suitable for these purposes. The two most common types are proficiency tests and achievement tests. If the purpose is to inform a future school or an employer about the level of proficiency, a proficiency test is appropriate. An achievement test, on the other hand, is meant to check how much the test taker has learned from the course lectures. This is useful if the

purpose is to give the students feedback on what they have learned from a course. It may also contain elements of language proficiency, but it is more specifically connected to a certain grammar and vocabulary of a given course syllabus. In order to get a high score, you may not need an overall proficiency, but it is sufficient to master the topics of the course. Two other purposes mentioned by Simensen (2007) are informing teachers and students about problem areas, and selecting students for courses with limited enrolment. In the first case, a diagnostic test may be used, and in the last case an aptitude test. The two last purposes are more peripheral, but she mentions cost-effectiveness assessment in order to inform school authorities of how the results are compared to resources, and tests for the purpose of providing data for research. Proficiency tests and achievement tests, as the most common forms, are also the most relevant for the present thesis.

2.2.3 Language test design

In the following the basic assessment types are discussed, followed by a closer look on types of language tests, followed in the next section by the choices that need to be made when developing language tests. Regardless of the test types mentioned above, there are several ways to design the tasks in a language test. The table below shows a simple overview based on a framework by Bachman and Palmer (2010):

Table 2: Language Task Characteristics

Aspect	Characteristics
Rubric	Instructions (specifications) Structure (sequence, weighting, time allotment) Scoring procedure (record, criteria, raters)
Input	Form (oral/visual, language/non-language, native/target language) Type (item, prompt, input for interpretation)
Expected response	Form (oral/visual, language/non-language, native/target language) Type (extended, limited, selected)
Input vs. Expected response	Scope (broad, narrow) Directness (direct, indirect)

Rubric

As the table above shows, there are four main aspects of a test according to this framework. The test rubric includes characteristics dealing with the organization and purpose of the test. It deals with the instructions, time allotment, structure and scoring procedure. The instructions

may vary in terms of language and explicitness. The time and structure characteristics concern the sequence and division of tasks in a test. And the scoring procedure is a characteristic of many variations. There is a choice of how to record the assessment result, as a score or as a description or both. Moreover, there are variations in criteria for correctness and choices concerning raters. The criteria will specify what should be scored, as well as how the levels of language should correspond with the scoring. Rating scales may also vary greatly, for instance several scales for different measures, or one scale being the sum of the scores on each task. The rating itself may be done by a single rater or multiple raters, and depending on the test, by human raters or computers.

Input and expected response

Two other aspects are input and expected response. What expected response means is self-explanatory, and input is everything that needs to be processed in order to give a response. The input may be simply a headline for a writing composition, or it may be a long text to read and comprehend before answering a task. Similarly, the expected response may be to tick off a multiple-choice item, or to produce a lengthy written text. Moreover, both of these aspects may vary in terms of whether they are aural or visual, language or non-language, and in what language, native or target language. But more importantly is perhaps the type of input and the type of the expected response, and there are three types of each. Two of the input types are item and prompt. A prompt is a directive which is intended to elicit an extended production response, while an item is input intended to elicit either a selected or a limited production response. Input for interpretation is the type which is not meant to elicit a response in itself, but is rather language meant to be the basis for a response. This is everything but the instructions and the items or prompts. It is typically a text that the test taker might be asked to analyze. The three types of responses are those mentioned in connection with the input and they are fairly self-explanatory. A selected response is typically a multiple-choice selection. A limited response consists of one word or a phrase, and an extended production response is anything longer than that.

Relationship between input and expected response

The relationship between input and expected response is the last area, which may vary in terms of scope and directness. A broad scope means it requires processing of global issues in

the input, while narrow scope may only require comprehension of certain details in the input. The directness relates to whether the information in the test itself is sufficient for a successful response. A direct task will not require other information than what is given in the test, while an indirect one will (Bachman & Palmer, 2010).

This short glimpse into the framework of language task characteristics shows the many options available for test developers in designing a language test. All of these characteristics may influence the outcome and the validity, and especially those related to the scoring procedure. This framework will be referred to in the test descriptions in chapter 6.

2.3 Defining the construct

The planning phase prior to an assessment was mentioned above. For formal assessment, this is called test development. In test development, the test construct is a fundamental part. Simply stated, a construct “is a concept (or a group of concepts) used to explain what the test tests” (Luoma, 2001, p. 4). Like a measuring tool measures physical attributes like distance and heat, a test tries to measure people's inner abilities. A thermometer measures heat; a language test measures language ability. In this case, language ability is the construct since that is what is tested. Heat is relatively easy to measure accurately because it is a physical attribute, and there is scientific agreement on what heat is. Language ability, however, is not anything physical, nor is there one single scientific definition. The ideas of what language and language abilities are continuously develop and change according to new research. A construct is such an idea. It is not something objective, or given by nature, such as heat, but rather a human theory or construction, hence the name. For a given language test, the construct may for example be language proficiency, or listening comprehension. The following extract from Alderson and Banerjee's review article on construct validity underlines importance of constructs:

An emphasis on the centrality of constructs – what we are trying to measure – requires testers to consider what is known about language knowledge and ability, and ability to use the language. Language testing involves not only the psychometric and technical skills required to construct and analyze a test but also knowledge about language: testers need to be applied linguists, aware of the latest and most accepted theories of language description, of language acquisition and language use. They also need to

know how these can be operationalised: how they can be turned into ways of eliciting a person's language and language use (Alderson & Banerjee, 2002, p. 80).

As this article points out, there are three levels of abstraction involved. The most abstract level is the theoretical construct, as described by the latest scientific research. This is a comprehensive scientific theory of the concept, like, for instance, a complete theory of writing ability, as the next chapter will present. The theoretical description forms the basis for the next level, which is the construct definition used in a test. The construct definition is meant to be readable for everyone involved in the test, and must therefore be more specific and concrete. It is a description of the abilities that the test should measure, and the degree of details involved may vary. Finally, the construct definition is operationalized, and we have the most concrete level, which is the actual test. The test is therefore based on the construct definition, and should give the test takers a chance to show the abilities that the construct definition describes.

2.4 Defining construct validity

Construct validity is considered the main concern in assessment by measurement theorists (Luoma, 2001, p. 72), but the view on how to define it has gone through a thorough evolution the last decades. The current view may be a complex concept to grasp, but starting with an outline of the traditional and simpler view may help to unpack it. For that reason, this section will start by defining construct validity in more simple terms, and the reasons why the simple definition fails will then lead us to the presentation of the current view. The last subsection explains the process of validation, that is, how to obtain construct validity.

2.4.1 The simple and outdated view

The simple definition can be explained as the ability of a test to measure what it is meant to measure. Since the construct is what a test is meant to measure, one may say the ability to measure its construct. This is why it is called construct validity. Such a definition, although it is not fully accurate, is found in literature even from this decade (Simensen 2007, p.253). But it is also found in a definition from 1937, stating that construct validity is how well a test “measures what it purports to measure” (Garett 1937 as cited in Luoma, 2001, p.324). What this means is easiest explained when contrasted to the notion of reliability.

Validity is a matter of whether a test measures what it should, while reliability is a matter of measuring accurately or consistently. If a test does not measure accurately enough, it is not, in fact, measuring anything at all. If a thermometer is unreliable, and sometimes showing ten degrees extra, sometimes not, is it then really measuring heat? So if you take away the reliability completely, there is no validity left either. But a test may have complete reliability, and still have no validity. This is the case if it measures accurately something other than what it is meant to. This is comparable to a thermometer accurately measuring air pressure. It does not help that the air pressure measurements are accurate if you are interested in the temperature of the room. In other words, a test that lacks either of these two qualities is useless. In contrast, construct validity, in the simple sense, is the quality of whether a test measures what it is meant to measure or not.

2.4.2 The broad and current view

Although this simple definition gives us a notion of what the validity concept is about, it fails on certain points. The reason why the definition above fails is apparent from the following quote by Samuel Messick:

Validity is not a property of the test or assessment as such, but rather of the meaning of the test scores. Hence, what is to be validated is not the test or observation device per se but rather the inferences derived from test scores or other indicators - inferences about score meaning or interpretation and about the implications for action that the interpretation entails (Messick, 1996, p. 245).

So if a test of French grammar is given to a class in German literature, there is nothing invalid about the test. The test is meant to test French grammar, and it measures French grammar. But if one would make interpretations on the basis of the result of that test, and say something about the students' knowledge of German literature, there is a validity problem. Such inferences would be invalid. Validity is therefore not whether a test measures what it is supposed to measure, but “the degree to which evidence and theory support the interpretations of test scores entailed by specific uses of tests” (AERA, 1999 as cited in Luoma, 2001, p. 62), or more simply: “the extent to which test interpretations and uses can be justified” (Chapelle, 1999, p. 258).

This definition differs from the traditional one on two vital points. Firstly, as stated above, the object of validation is the interpretation and use of test scores, and not the test or the test scores themselves. Thus a test may be valid in the traditional sense, but it may be used in an invalid way, like the French grammar test mentioned above. The definition above speaks about justifying interpretations and uses of the test. Interpretation of a test is what we have defined as the second step of the assessment process, and the use of a test is the decisions based on these interpretations. The first difference is therefore that construct validity does not refer to the quality of a test, but to the interpretations and use of it. Secondly, it is a matter of justification rather than qualification. It is not really an objective quality that can be proved beyond all doubt. It is about how justified the interpretations and uses are. But how does this justification happen? This is what validation is about, and will be discussed below.

Today, construct validity is considered an overarching concept, of which aspects of content and criteria are integrated parts. Even reliability may be viewed as an integral part of construct validity rather than in tension with it (Luoma, 2001, p.7). By taking a brief look at how current validation is done, this broad concept will be clearer.

2.4.3 Validation

Since the idea of construct validity has changed drastically the last decades, the validation process is also different. Traditionally, it has been sufficient to provide a correlation with another valid measurement to prove that a test is valid. As mentioned above, validity is not the property of a test, in fact, one cannot prove complete validity. One definition from the 1999 Standard states that “the process of validation involves accumulating evidence to provide a sound scientific basis for the proposed score interpretations” (as cited in Luoma, 2001, p. 74). Thus instead of providing certain proof, it is about justifying the use of the score interpretations through scientific evidence. But two crucial questions arise. How do we go about accumulating such evidence? And how much evidence is enough?

The answer to the second question is given by Bachman and Palmer (2010) as they speak about accountability. They say that because assessment will affect people's lives, we “need to be able to justify the uses - consequences and decisions - of a particular assessment so that we

can be accountable to ourselves and to other stakeholders” (Bachman & Palmer, 2010, p. 92). It is a matter of accountability, so it is enough evidence if the stakeholders are convinced. In this sense, validation is convincing the people most likely and most directly affected by its use and consequences that the use of the assessment is justified. The affected ones are most directly the test takers, but may also be fellow teachers, parents or school administrators.

Answering the first question about how to accumulate the evidence is more complex. Bachman and Palmer suggest to first articulate the claims about the test use, and then to try to find evidence to back up the claims. But there are different types of evidence, and several claims that need evidence support. Luoma states that

All the systematic influences that can affect scores should be investigated in a validation exercise. Thus the scope of validation inquiry is quite broad. In addition to the test scores themselves, the test, the testing procedures, the context in which the test is implemented, and the processes that the test takers and assessors go through during the testing process must be investigated to explain the meaning of the scores (Luoma, 2001, p. 72).

To get an overview of the systematic influences that can affect test scores, it is useful to look at the three phases related to assessment. The first phase is test development, the second is the assessment process and the third is the test use. Figure 1 below illustrates the three phases,

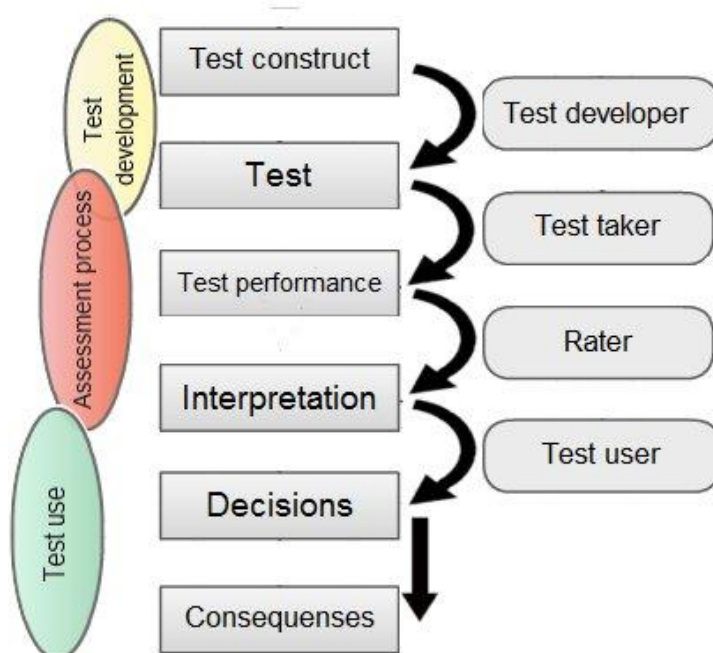


Figure 1: Validation process of a test

and the steps that occur in and between each phase. Each of these steps is an inference and represents a claim that needs to be supported by evidence.

The first step, as shown in the figure, is operationalization, which is the process of creating the test based on a test construct. In a validation process, one would have to provide evidence that the test really reflects the construct. This step is the most relevant for the present study, and section 5.4 describes potential errors that may occur here. Next is the test performance, which will depend on the tasks in the test and how the test taker responds to them. The interpretation step is largely dependent on the rater or raters of the test. Ensuring that the raters interpret the test performance according to the construct and that they score them reliably is also an important part of the validation process. Again, this is particularly relevant, since I include an analysis of reliability in this study. The final phase is test use where decisions will be made based on the scores, and decisions lead to consequences. If the stakeholders of a test can be convinced that all these steps are justified to make, the validation process is successful. On the other hand, if there is evidence showing otherwise, changes should be made.

To sum up, validation is not proving validity as a test quality, but convincing those affected by the assessment that the decisions made on the basis of assessment interpretations are justified. Secondly, this process happens by articulating arguments about the inferences that the assessment assumes, and providing evidence for them. Thirdly, there are several inferences assumed in the use of assessment scores to make decisions, and many possible influences, all of which must be taken into account in a full validation process.

3 Writing

If a teacher conducts dictation in class, the teacher is speaking, while obedient students are *writing*. If one of the less obedient students receives an SMS during the dictation, and replies, he is also *writing*. If the brightest student ten years later becomes a textbook author, what he does for a living is *writing*. Clearly, these three activities are really different although they all may be called writing. In dictation, the students only need to know the correct spelling and how to write the letters. Typing an SMS requires reading the message, and finding out what to respond, but it often requires less concern of correct spelling and grammar. The textbook author, however, needs to consider all the linguistic features as well as the content. He might not even have all the necessary knowledge for writing the book before he starts. How can such differences in writing activities be explained? Or perhaps, how can there be a single definition of what writing is when there is so much variation?

Bachman and Palmer (2010) suggest avoiding speaking of writing as a single skill, and instead define each writing activity separately. Weigle similarly concludes that “Instead of attempting an all-encompassing definition, then, it may be more useful to begin by (...) the types of writing that are likely to be relevant” (Weigle, 2002, p. 4). Therefore, this chapter will begin by describing different types of writing, and consider the relevant types for writing English in upper secondary school. The next section will go on to define it as a cognitive process, and focus on the difference between composition by mature and immature writers. Building on the cognitive model, the last section will expand on the language component of writing.

3.1 Types of writing

In order to classify types of writing, we may speak of the three dimensions of the process. The three dimensions are audience, intention and level of cognitive processing. The following outline is based on a model of writing discourse from Vähäpääsi (1982, as cited in Weigle 2002), and shows the variation of writing activities and their basic differences.

By the first dimension, audience, one means the ones for whom the text is written. Is it written primarily for oneself, or for others? Texts are not always meant to be read by others than the writer himself. This is usually the case for lecture notes, shopping lists or diaries, for example. Most other types of texts, such as articles or letters, are written with some other audience in mind. Writing for oneself is a common everyday activity, and does not require the same accuracy as when writing for others. Next, by intention one means the purpose of the text. What do we mean to achieve with our writing? Vähäpääsi (1982 as cited in Weigle 2002) differentiates between the following six intentions:

1. To learn
2. To convey emotions
3. To inform
4. To persuade
5. To entertain
6. To stay in touch

If the primary audience is oneself, the dominant intention is limited to the two first points on the list. Writing for oneself is commonly done in order to learn or remember something. It is also possible, though, to convey emotions to oneself when writing a diary. Otherwise, if the dominant intention is any of the other options, the text must be meant for an audience. If you write an email or a postcard, the audience is the addressee and the intention is to stay in touch. If you write an editorial, the audience is the readers of the newspaper and the intention is to persuade. Thus these purposes make for different types of writing that each requires different skills. Although one may argue that it is easier to convey emotions than to write persuasively, these categories are not meant to be ranked in any particular order. Moreover, the categories are not mutually exclusive. A letter to a friend may both be to keep contact and to inform of your health condition. Thus a text may serve several purposes, but it may still be categorized by the most dominant of these.

With the two dimensions mentioned this far, we can describe some of the differences between dictation, SMS-typing and textbook work. Dictation is primarily for oneself, while the other two are to be read by someone else. Moreover, they all differ in terms of dominant intention. The purpose of dictation is to learn something; it may be used to learn the spelling of a language. An SMS, however, is written to stay in touch. And the objective of a textbook is to give information about a subject matter. However, there is also a difference at another level

which these dimensions do not account for. If you subscribe to a newspaper and are asked to fill out your name and address in a form, the intention of the writing is to inform, and the audience is others. Even though the intention and audience are the same as in textbook production, it is clearly not the same type of writing. This difference is therefore explained by the third dimension. The third dimension, cognitive processing, is a taxonomy of three levels involving an increasing degree of difficulty.

1. Reproduce
 2. Organize / reorganize
 3. Invent / generate
- (Vähäpääsi, 1982 as cited in Weigle, 2002, p. 8)

The most basic level is reproduction. This is the case in dictation, or when filling out a form. All it requires is the skill to form the words correctly. It simply consists of writing down a given text as it is. The next level, called organization, or reorganization, is more demanding. A typical example of writing at this level is making a summary of a story. In a summary, the main points must be outlined, and the details left out. It requires the ability to reorganize known information. Everything that needs to be written down is still known to the writer beforehand. The most demanding level is to invent or generate written language. This is also distinctly different from level 1 and 2. In the basic levels, the writer writes from the knowledge he has. At level 3, the knowledge of the writer is not only informing the writing process, but also being informed by it. All the information needed for the writing task is not known to the writer beforehand. It means that the knowledge of the writer changes as he or she writes. A writer may make up his mind on a subject during the process. This is the type of writing that characterizes the work of a textbook writer. The cognitive model in the next section (see 3.2.2) will further illustrate the difference between level 3 and the other levels, also known as knowledge telling and knowledge transformation.

In our search for a definition of writing ability, we have looked at the different types of texts that people write. Weigle (2002) indicates that the need for these different types of writing varies with the language user. The situation of the language user may be such that writing is hardly needed at all in the real world. Being able to keep up a conversation may sometimes be enough. This may often be true for learners of foreign languages. However, “For students

nearing the end of compulsory education and intending to go on to higher education, Type III³ writing takes on greater importance” (Weigle, 2002, p. 11). The learners that are relevant for this paper are nearing higher education, which makes level 3 writing highly relevant.

Although it is not a matter of survival, it is a matter of importance for academic skills. The fact that the foreign language is English makes the need for these skills even greater (Weigle, 2002, p. 7).

3.2 Cognitive model

This section will provide a cognitive model of writing. The model is a description of the mental activities that go into composing written texts. It is based on the models of writing by Hayes (1996 as cited in Weigle, 2002), and Bereiter and Scardamalia (1987). In both of these models there is mention of input, a process and output. The output is obviously written composition, but what is the input and what occurs in the process?

3.2.1 Input and process

The models operate with internal and external inputs. The first external input is the writing assignment. It may be a school or work assignment, but also simple everyday writing activities. The assignment will normally specify topic and genre. Secondly, once some of the text has been written down, that text will also serve as input for further writing. Reading through your own text may give clues on how to continue. In addition to these two inputs, which Hayes terms the physical task-environment, there is also an external social environment. There may be a co-writer, and there may be an actual audience which both may contribute to the external input.

The internal input, on the other hand, is primarily provided by what is known as the long-term memory. Our long-term memory is where all our knowledge is stored, and this knowledge serves as input in the writing process. The two major parts of this knowledge are called topical knowledge and language knowledge, both of which are indispensable in order to write. Topical knowledge, also called content knowledge, is a term for the knowledge of the world around us. This knowledge is needed in order to have something to write about. Language

3 The same category as I have called level 3 above; that is to invent/generate

knowledge is knowing *how* to write about the different topics. “Language knowledge can be thought of as a domain of information in memory that is available to the language user for creating and interpreting discourse in language use” (Bachman 2010, p.44). Hayes also mentions genre knowledge, audience knowledge and task schemas as part of the long-term memory storage. However, as the subsequent section will explain, these are all integral parts of the current concept of language knowledge. The other internal factor is motivation, which also strongly influences the writing process. This means that the writer's goals and attitudes will determine the pace and the effort involved in the process. Although motivation is not irrelevant in connection with writing assessment, it is beyond the scope of this thesis to go into motivation theory. To summarize it simply, the assignment and the text so far constitute the two main internal inputs, whereas topical knowledge and language knowledge constitute the external.

All the input above is processed in order to produce the written composition. The processing is carried out by our working memory. The working memory, in contrast to the long-term memory, does not store knowledge over time. It only brings forth the information needed in the moment, and has a limited working capacity. How does this process look? According to Bereiter and Scardamalia's (1987) research, this process may take two different forms, either a simple form, or a complicated one. The simple form is called knowledge telling, and describes the basic reproduction and organization levels presented in section 3.1 above. The more complex process is called knowledge transformation, and corresponds to level 3 in the same model, to invent or generate.

3.2.2 Knowledge telling and knowledge transformation

The process of knowledge telling is quite simple, and resembles in some ways the process of speaking. Speaking usually happens unplanned where thoughts are directly translated into oral language. What comes out of the mouth is part of the storage of knowledge inside. Similarly, knowledge telling is writing down the ideas that come to mind from our memory. The ideas come from the input factors mentioned above. The topic of the assignment may trigger some knowledge from memory, and produce potential ideas. In the same way, the text written so far, may elicit new ideas about how to develop the text further. The first step is therefore to produce ideas from the available input. Secondly, these ideas are subject to a test of

appropriateness. Is this idea suitable for the assignment, and will it support the idea of the text? If the answer is positive, the idea is written down. If not, one returns to the input for another idea. When the writer is out of ideas, the writing process is over. It is a simple and natural process that “requires no significantly greater amount of planning and goal setting than does ordinary conversation” (Bereiter & Scardamalia, 1987, p.9).

What makes knowledge transformation different is the generation of new knowledge through the writing process. Writing is not only a means of expressing knowledge, but also a means of learning or gaining new knowledge, requiring a differently structured model. The starting point in this case is a problem analysis and goal setting based on the assignment in question. The problem analysis is worked out in two areas, the content problem space and the rhetorical problem space. This means that there are two types of problems that the writer needs to work with. One is related to belief and content. This requires use of topical knowledge from our memory, or from outside sources. The other deals with rhetorical issues, how to complete the assignment, and requires use of language knowledge. Thus there is a problem-solving process going on before the ideas are even ready for the test of appropriateness. It is a process of working out issues related to content and language, where the knowledge is not only contributing to solve the problems, but is also affected by it. Writers using this model “consider not only changes in the text but also changes in what they want to say. Thus it is that writing can play a role in the development of their knowledge” (Bereiter & Scardamalia 1987, p.11). There is a two-way interaction between the knowledge and the problem spaces. After finding solutions to the problems of content and language, they become input for the knowledge telling process as described above. Instead of writing down ideas that come directly from the assignment input, as in knowledge telling, thoughts are reworked and reconsidered in a thorough process before they are written down (Bereiter & Scardamalia, 1987). While knowledge telling is described as a natural process, this is described as hard work.

3.3 Communicative competence

Although the cognitive model has shown that writing requires more than knowledge of language, it still remains a fundamental component. This last section will define what is

meant by the term language knowledge in the model above. The Common European Framework of Reference for Languages (CEFR) uses the term communicative competence (Council of Europe, 2001) to describe language. Since language is primarily a tool of communication, knowing a language is equal to knowing how to communicate in that language. The following framework of communicative competence is based on the most recent frameworks of the CEFR and Bachman and Palmer (2010).

The term communicative competence was first coined by Dell Hymes. He defined it as knowing “when to speak, when not, and as to what to talk about with whom, when, where, in what manner” (Hymes, 1972 as cited in Simensen, 2007, p. 72). In other words, knowing a language is more than just knowing grammar and vocabulary, and mastering a language is more than being able to translate text correctly. To account for all the aspects of language knowledge, the CEFR operates with the following three main components: linguistic competence, socio-linguistic competence and pragmatic competence. Before presenting these components in more detail, it is useful to notice that this is a fairly comprehensive model of language knowledge. Many language assessments “will focus on assessing only one or a few of these areas of language knowledge” (Bachman & Palmer 2010, p.44). However, “The design of every language assessment no matter how narrow its focus, needs to be informed by a broad view of language knowledge” (Bachman & Palmer 2010, p.44).

3.3.1 Linguistic competence

Bachman and Palmer (2010) divide this competence into a grammatical and a textual part. The grammatical part deals with word and sentence level issues, while the textual part covers the parts of discourse that are above the sentence level. A traditional view of language has always considered the grammatical features. It includes vocabulary, syntax, morphology, orthography and so on. It concerns the rules of word formation as well as word combinations and fixed expressions. It contains rules for how the smallest parts of language can be combined to form words and sentences. If all the grammatical rules are followed, there will be no spelling or grammar mistakes. It is not sufficient, however, for creating a coherent text. The textual aspect of linguistic competence is what makes a text a text, rather than unrelated sentences next to each other. In order for sentences to create a coherent whole, they must be in relation to each other, and such relations are created by using cohesion. “Cohesion refers to

the linguistic devices by which the speaker can signal the experiential and interpersonal coherence of the text” (Thompson, 2004, p. 197). Some examples of such linguistic devices include expressions like “however”, “in order to” and “because”. These words are used to bind the sentences together logically, but there are also other ways to create coherence. Use of words related to the same field creates the common topic of a text, and use of pronouns makes the references clear from one sentence to another. One may also speak of organizing the text at an even higher level. This means placing the passages in a sequence according to the conventions of the genre, whether it is an argument, a narrative or poetry.

3.3.2 Socio-linguistic competence

Linguistic competence will enable you to write grammatically correct and coherent language. Socio-linguistic competence adds the aspect of choosing the grammatical structure or vocabulary fitting for the setting. Since language is used for communication, it is always used in a social context. This competence is the ability to relate language to that context according to the norms of the language and society. The choice between saying “I’m good” or “No, thank you. I am perfectly satisfied” when offered more food at the table is an example of this. Both options are linguistically acceptable, but the social context will determine which one is more correct or suitable. The social context consists of the setting and the people involved, and the relationship between them. Having socio-linguistic competence means being able to consider social relations, politeness, level of formality, dialects and also cultural references in a language situation. Without this competence, communication will not always succeed.

3.3.3 Pragmatic competence

Imagine a group of good friends eating pizza together and watching ice-hockey. The host asks one of his friends if he wants more pizza, and the friend answers: “No, thank you. I am perfectly satisfied.” This seems correct linguistically, but quite odd socially, in accordance with the socio-linguistic competence described above. “I’m good” would seem to be a more fitting answer. However, it may be that this is not due to lack of socio-linguistic competence, but rather conscious use of pragmatic competence. The reply may be acceptable if it serves some other function. Pragmatic competence deals with the functions of language use. In this case, one could picture that this was meant to be funny. The fact that the reply breaks with the social setting makes it humorous. Understanding what someone means is not equivalent to

understanding the grammar and the vocabulary. It is not even always enough to understand idioms, dialects and cultural references. It requires pragmatic competence, the ability to relate text to the intentions of the language user (Bachman & Palmer 2010, p.46). It is the ability to interpret the underlying meaning of an utterance. When meeting a pregnant woman nearing her due date, people might ask: “Hasn't the baby come out yet?” The form of the sentence is an interrogative form, but it is not intended as a question in the normal sense. The utterer already knows the answer, and is only interested in small-talk, and perhaps the function was a conversation starter. Being able to create as well as understand these intended meanings is what is meant by pragmatic competence.

3.4 Summing up

To sum up, there are different types of writing activities, and which activities are useful depends on the language learner and his or her situation. For many it is enough to be able to keep a conversation, and writing is hardly needed at all. For others, especially those in higher education or nearing it, it is useful to master writing of all types, and even writing as knowledge transformation. The cognitive model by Hayes (1996 as cited in Weigle 2002) asserted that writing consists of handling several input sources, and processing them into coherent text. In an attempt to summarize the model more simply, I have come up with the following list of skills required for writing:

- understand the instruction of how and what to write (*external input*)
- know what to write about (*internal input: topical knowledge*)
- know how to write it... (*internal input: language knowledge*)
 - grammatically correct (*linguistic competence*)
 - textually coherent (*linguistic competence*)
 - situation appropriate (*socio-linguistic competence*)
 - genre-specific (*socio-linguistic competence*)
 - according to the intention of the instruction (*pragmatic competence*)

The way in which these skills are applied to produce a written text may differ. Immature writers will often rely on knowledge telling, which is simply to write what is known and seems appropriate according to the instruction. More mature writers will apply knowledge transformation where the writing is produced in a demanding process where learning also takes place. The next chapter will present how writing is defined in the national curricula of Norway and Finland, starting by presenting their different educational systems.

4 Norway vs. Finland

In the national examinations examined in the present thesis, it is the national curricula that define the test constructs. This chapter will therefore provide the contexts in which these examinations exist by comparing the school, curricula and examination systems of the two countries. It goes on to describe the English syllabi for upper secondary school, and finally to present the writing constructs as the syllabi define them.

4.1 A comparison of the school and examination systems

In general, the school systems of Norway and Finland are quite similar. In both countries, compulsory education continues until the age of sixteen. This is followed by a choice between general or vocational upper secondary school, which gives access to different types of higher education. It is also the case in both countries that compulsory school is divided in two, starting with primary school, and ending with three years of lower secondary school. A small difference in this structure concerns the age for starting school. Norwegians start school at the age of six, and have ten years of basic education, while the Finns start at seven and have nine years of basic education. There are, however, a couple of other differences that are relevant to mention. This section will briefly describe some aspects from each nation in the following four areas: national curricula, organization of foreign language teaching, courses in upper secondary school and examinations.

4.1.1 National curricula

The teaching in upper secondary school is in both countries regulated by a national document. The Norwegian national curriculum is from 2006 and is called the Knowledge Promotion (LK06). It is a common curriculum for all of primary and secondary school. It comprises The Core Curriculum and the Quality Framework, Subject Curricula and Distribution of teaching hours per subject. The first part describes the school and its role beyond the mere subject related issues such as human values and identity. The Subject Curricula, or what one may rather call the subject syllabi, deal with the purpose and competence aims for every school subject. This is all decided on a national level and regulates the practice for every Norwegian school as the Education Act of 1998 specifies:

The Ministry issues regulations concerning levels and programme areas, concerning subjects, educational objectives, the scope and content of instruction in subjects and the organisation of the education and training. (...) The teaching staff is to organise and carry out their teaching in accordance with subject curricula issued pursuant to the present Act (The Education Act, 1998, para. 3–4).

In the Finnish system, the national curriculum is called the National Core Curriculum for General Upper Secondary Schools (NCC), and is clearly separate from the curriculum of basic education. It covers the values initially, and has a chapter for every subject where the main objectives and courses are described. In contrast to the Norwegian document, it does not contain detailed lists of specific competence aims for every level. In fact, NCC is meant to be the basis on which local curricula will be drawn up by the local education provider. The curriculum states that “Education providers may decide how to draw up their curricula on the basis of the National Core Curriculum” (Finnish National Board of Education, 2004). Two of the areas in which the local education provider has some freedom are language programs and specialization courses. The two following sections will elaborate on these issues.

4.1.2 English in primary and lower secondary school

In Norway, English is a compulsory subject through all of basic education, and it is listed in the curriculum as a subject separate from other foreign languages. The Norwegian curriculum also prescribes 138 hours of English for the first four years of school, and after two years there are a set of competence aims to be attained. Even though it does not specify the amount of hours to use for English each of those two first years, it mentions English as a subject called ENG0001, which is intended for first grade (Udir, n.d.). It is slightly more complicated to explain when English teaching starts in Finland, since it may vary from school to school. English is, in fact, only described as one of many foreign languages to choose from, and does not have a separate syllabus as in Norway. And since there are two official languages in Finland, Finnish and Swedish, learning the second national language is also prioritized. But the following diagram from the Finnish National Core Curriculum for Upper Secondary School 2003 will illustrate the issue of foreign language courses.

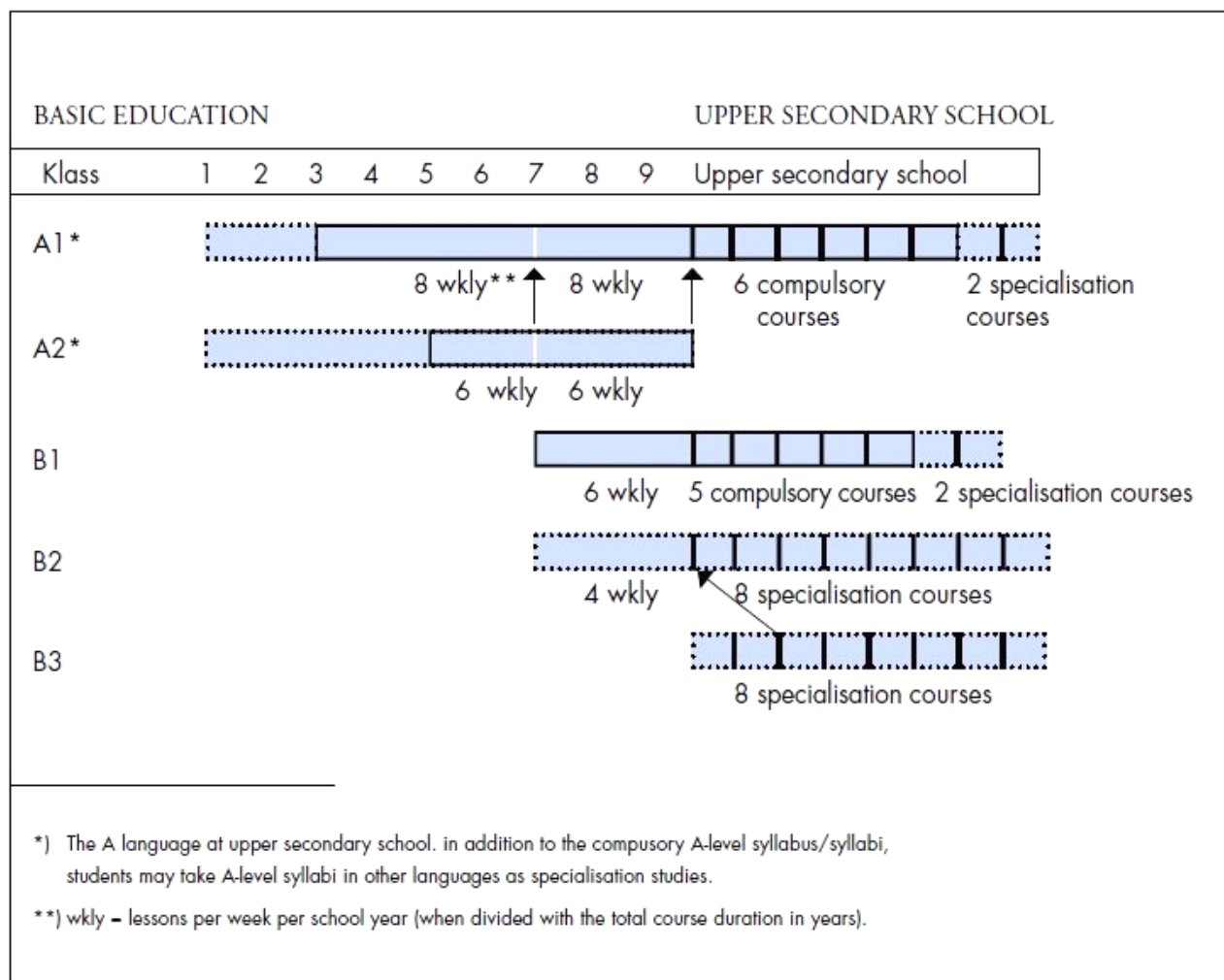


Figure 2: Distribution of language levels in Finnish compulsory education, From Finnish National Board of Examination (2004)

As figure 2 shows, they operate with five potential foreign languages. They are labeled A languages if they are started in elementary school, and labeled B if they are started later. Since each school in Finland needs to develop their own local curriculum based on the national one, it is also up to every single school to decide which languages they offer at each of these stages, and there seems to be no regulations in the national documents about when English should begin. This means that, theoretically, it is possible to complete nine years of compulsory school in Finland without having studied English. And even if English is chosen as the first foreign language, A1, it is not until third grade that it usually begins. But as the dotted lines show, it is possible to start earlier, and in 2009, 8 percent of pupils began the A1 language studies in the first grade and those who started in second grade accounted for 14 percent. In the same year, more than 90 percent of all the pupils in Finland chose English as their A1-level language (Suomen kieltenopettajien liitto, n.d.). So, although it is theoretically possible not to choose English in the Finnish school, the actual situation is not so far from the

Norwegian one after all. English is considered an important language in both nations, and is taught from the very first years of elementary school.

4.1.3 Courses in upper secondary school:

This passage will describe perhaps the greatest difference between the two school systems at upper secondary level, namely the organization of courses. Since the examinations in question are not relevant for students in vocational upper secondary school, I will only describe the situation for general upper secondary school.

In Norway, there are three levels in upper secondary school that normally last for one year each, Vg1, Vg2 and Vg3. The duration of the courses in each level is also one year, and some of them are compulsory, others elective. During the first year (Vg1), there are no elective studies offered, but eight compulsory courses that must be completed. When starting in Vg2, it is required to specialize in a certain field, such as the natural sciences or languages, and choose a combination of elective subjects within that field. This specialization program will continue for two years, and in Vg2 and Vg3, there are five and four compulsory courses in addition to the program. In order to graduate, all three levels must be completed, comprising a total number of 2523 lesson hours (Udir, 2011a).

The Finnish situation is different in this matter because the courses are shorter and more independent of one another. Instead of yearlong courses, there are courses that comprise only 38 hours⁴ of instruction. Similar to the Norwegian system, there are compulsory as well as specialized, elective courses, and the education is completed when a minimum number of 75 courses are attended and approved, whereof at least ten must be specialization courses. Within a certain framework, and normally within a time frame of three years, the students are free to set up these courses in an individual study plan. The education may also be completed in shorter or longer time, but not longer than four years without special reasons. (Westermarck, 2011).

4 On average

This difference obviously affects the English language courses as well. In Norway, there are altogether four English courses in upper secondary school. One of them is a compulsory course in Vg1, and the remaining three belong to specialization courses for those who choose English in their study program. In Vg2, the specialization course is called *International English*, whereas Vg3 offers two options for those who have passed the Vg2 course: *Social Studies English* and *English Literature and Culture*.

In Finland, the schools offer six compulsory and at least two elective specialization courses in English. These eight courses are specified by the national curriculum, and further description of their content is presented below in section 4.3. The local schools may also choose to offer additional specialization courses in English, according to their resources. Because of the individual study plans, the students may each decide when to attend these courses during the three years of school.

4.1.4 Examinations:

Finally, it is worth mentioning the differences in how the national examinations take place. In Norway, everyone is to have one oral and one written examination at the end of lower secondary school, as selected by the education provider (Udir, 2009). In Finland there are no national examinations at this level. In fact, the only nationally given examinations in Finland take place during the last year of upper secondary school, and are organized in the following manner, as explained by The Matriculation Examination Board in Finland:

The examination consists of at least four tests; one of them, the test in the candidate's mother tongue, is compulsory for all candidates. The candidate then chooses three other compulsory tests from among the following four tests: the test in the second national language, a foreign language test, the mathematics test, and one test in the general studies battery of tests (sciences and humanities). As part of his or her examination, the candidate may additionally include one or more optional tests (Ylioppilastutkintolautakunta, n.d.).

In other words, the students sign up for the tests that they choose within a given framework. In Norway, there are examinations after every year of upper secondary school, and the selections made by the school authorities. In Vg1, 20 per cent of the students are selected for an oral or written examination, and in Vg2, all students are to be selected. In the third and last

year, Vg3, everyone is to take the compulsory examination in Norwegian, and two additional written, as well as one oral examination in the other subjects (Udir, 2009). This means that, in Norway, those who are selected for the English examination in Vg3 have completed three yearlong courses of English. In Finland, anyone who has completed the six compulsory courses can sign up for the English matriculation examination.

4.2 The English syllabi

The following sections will elaborate on the Norwegian and Finnish English syllabi for upper secondary school. After a more general description of the syllabi, the last subsection will outline the writing constructs from these syllabi, as they are relevant to the examinations in question.

4.2.1 The Norwegian English syllabus

As mentioned above, there is a compulsory English course in Vg1, but since this study focus on the Vg3 examination, the syllabus for the specialized study program is presented here. Due to reasons explained in the next chapter, we will also focus mostly on the areas related to “English Literature and Culture”, hereafter also referred to as the literature course. Three sections of the Norwegian English syllabus are elaborated on below: the objectives of the subject, the main subject areas, and the competence aims within each of those areas.

Objectives of the subject

The first section in the syllabus presents the higher objectives of this English subject. The passage describing the objectives is an ambitious text with references to many skills and uses of the English language. In the first half, it emphasizes the role of English in a global society, and the need to know English in order to access vital information. Thus, an important objective is to help the learners understand global issues, and therefore be able to participate in social and working life. This knowledge as well as the language skills may open doors to other cultures, and the history of the English-speaking countries helps us understand the present situation of great English influence. The second half mentions the value of literary works, and how English literature, culture and language may contribute to personal growth, and help develop skills in critical thinking. It is also a goal that this subject may help the learners develop learning strategies for continuous learning and development.

Main subject areas and competence aims

While the objectives are fairly broad, and concerns all the three courses, this part of the syllabus tries to specify the content of each course. The content is divided in three so-called main subject areas, which are “Language and Language Learning”, “Communication” and “Culture, Society and Literature”. Every area and its scope are briefly described in the syllabus, and each area has a number of related competence aims. Next, I will present how it describes the main subject areas for the literature course, and most of the competence aims for the same course will be listed in the construct definitions at the end of this chapter.

The *Language and Language Learning* area mainly describes knowledge about linguistic features and terminology needed to analyze texts in relation to social and cultural issues. It is concerned with the development of tools to discuss the relationship between form and content of texts as well as other forms of expression. The aspect of learning strategies and the ability to evaluate own learning progress is also included here. The *Communication* area deals with the use of the language for communicating meaning. This is where linguistic features come in, such as nuanced vocabulary and coherence, but also socio-linguistic elements such as adapting language use to situations and genres. Most of the competence aims in this area are not about communication in general, however, but specifically related to literary and cultural issues. The third area, *Culture, Society and Literature*, includes more competence aims than any of the other two, and covers a large content area. It deals with interpretation, analysis and discussion of different types of literature, but it also covers cultural expressions of other types like film, music and architecture. Historical as well as current issues are covered, and Anglo-American culture as well as international culture.

Finally, it should also briefly be mentioned that in addition to these subject specific areas, there are four basic skills which also are part of the subject competence. This means that knowledge of these skills can be expected even if they are not stated in the competence aims. The skills include the ability to express oneself orally and in writing, the ability to read, numeracy and the ability to use digital tools.

4.2.2 The Finnish English syllabus

In the Finnish curricula, in the chapter on foreign languages, there are references to English as an A, B1, B2 or B3-level language. But since we already saw that it is most usual for English to be chosen as an A-level language, this is what we will describe here. The most relevant parts of the syllabus to examine are the objectives of instruction and the course descriptions.

Objectives of instruction

“The objective is for students to achieve the levels of the Language Proficiency Scale (...) as per the table below” (Finnish National Board of Education, 2004). The syllabus includes a Finnish application of the language proficiency scale from the CEFR. The table referred to is not included here, but it shows that the level to be achieved for A-level languages is B2 on that scale. The scale is divided in the four language skills: speaking, listening, reading and writing. For each of these skills it describes “can do”-sentences for a given level. Since the Finnish examinations in question only cover reading and writing, it may be enough to look only at those parts of the scale.

The *reading* description defines this level as being able to “read a few pages of text independently (...) about his/her own field or general topics” (Finnish National Board of Education, 2004, p.247) This may include texts of different genres, and texts that deal with different types of general subjects. At this level the learner is also able to identify general meaning and locate details in a long text, but idioms and allusions connected to cultural concepts may be hard to decode. The *writing* description mentions several relevant sub skills needed for text composition. In general, it states that at this level, the learner will be able to write a clear text of some detail, either personal or formal. It also mentions the ability to express views effectively and summarize information from different sources. The linguistic skills to be in command of at this level are broad vocabulary, demanding sentence structure, cohesive text production, orthography, grammar and punctuation. While errors in these do not cause misunderstandings at this level, there may be inaccuracies in style and expression.

In addition to achieving these levels of proficiency, the syllabus also includes four objectives that mainly deal with learning strategies and self-assessment. The first one, however,

highlights cultural sensitivity in communication, knowing how to express oneself characteristically of that language.

Course descriptions

There are eight English courses that all Finnish schools will need to offer for A-level language learners during upper secondary school. Two of them are so-called specialization courses, and six are compulsory. Each local school may also offer other relevant specialization courses in addition to these. The title of a course describes the main theme it deals with, and these themes are discussed from a Finnish perspective as well as the perspective of the target language culture. But other themes may be included in each course if they are of special interest to the students, or of special relevance to current issues. The courses emphasize different language features and skills, but all of them are to give the students the opportunity to speak, read, write and listen for different purposes, and each will pay attention to language structure, vocabulary and accuracy of language use. Use of authentic text or material will also be used, and attention paid to contrastive difference between English and one's mother tongue.

The three first courses are called: “Young people and their world”, “Communication and leisure” and “Study and work”, and cover such everyday topics. The language focus in the first one is on vocabulary, basic structures and expression of opinions. While the other two cover oral communication and strategies as well as writing skills and formal language. The three remaining compulsory courses are called: “Society and the surrounding world”, “Culture” and “Science, economy and technology”. These seem to be on a higher level with focus on speaking and reading at a more demanding level as well as writing texts for different purposes. The course on culture specifies that the students are to prepare a project on a chosen topic and present it. Finally, the two specialization courses are “Nature and sustainable development” and “Globalization and internationalization”. The focus on diverse development of language skills, and are more content related. The course on nature deals with using language related to nature and natural sciences, while the other one deals with global development and world views.

4.2.3 The construct definitions of writing

In order to do a systematic analysis of the selected examinations in relation to their construct, we need to extract a specific construct definition of writing from each country. The content of the syllabi described above is too complex and detailed. However, the competence aims of the Norwegian English syllabus and the Finnish “can do”-statements will function well in this matter. These parts of the syllabi describe a writing construct that can be used for this purpose. They sum up the essence of the construct, and are also quite specific measures. The points on self-assessment, as well as project related elements are excluded from this definition due to irrelevance for the present analysis. The relevant components for each writing construct are listed below, and listed first is the Finnish construct:

The English writing construct definition for the Finnish examinations:

- Can write clear and detailed texts about a variety of areas of personal interest and about familiar abstract topics.
- Can write routine factual messages and more formal social messages (reviews, business letters, instructions, applications, summaries).
- Can express information and views effectively in writing and comment on those of others.
- Can combine or summarise information from different sources in his/her own texts.
- Can use broad vocabulary and demanding sentence structures together with linguistic means to produce a clear, cohesive text. Flexibility of nuance and style is limited and there may be some jumps from one idea to another in a long contribution.
- Has a fairly good command of orthography, grammar and punctuation and errors do not lead to misunderstandings. Contributions may reveal mother tongue influences. Demanding structures and flexibility of expression and style cause problems (Finnish National Board of Education, 2004, p.247)

The Norwegian list is slightly longer, and is sectioned in three areas as explained above:

The Norwegian English writing construct definition:

Language and Language Learning

- elaborate on and discuss the relationship between form, content and stylistic register in sentences and texts
- have a command of the terminology needed for analysing works of fiction, films and other aesthetic forms of expression
- elaborate on and discuss distinctive linguistic features of texts from different genres, from different periods and regions

Communication

- use a nuanced, well-developed and precise vocabulary to communicate on literature and culture
- elaborate on and discuss lengthy and linguistically demanding discourses with general, specialized and literary content
- use suitable language, appropriate to the situation, in oral and written genres
- summarize, comment on and discuss differing viewpoints in fictional texts
- produce texts in a variety of genres with clear content, appropriate style, good structure, and usage that is precise and accurate

Culture, Society and Literature

- interpret a representative selection of texts from literary-historical periods in English literature, from the Renaissance up to the present time
- analyse at least two lengthy works of fiction
- analyse and assess a film and a selection of other artistic forms of expression within English-language culture
- interpret literary texts and other cultural expressions from a cultural-historical and social perspective
- elaborate on and discuss a selection of factual prose texts from English-language culture and social life
- elaborate on and discuss the cultural position of the United States and Great Britain in the world today, and the background for the same
- elaborate on and discuss current issues in international culture and the news media (Udir, 2006, pp.6-7).

The communication part of the Norwegian definition resembles the Finnish list in some aspects. Producing a variety of texts is clearly stated in both, and so are summarizing ideas, broad vocabulary and demanding linguistic structures. The Finnish definition is more explicit when it comes to stating grammatical features like orthography and punctuation, while situation appropriate language is only mentioned in the Norwegian one. However, the two other sections in the Norwegian construct, concerning cultural and literary topical knowledge, have no parallel in the Finnish one. From these descriptions, it is quite clear that the Finnish writing construct deals purely with writing proficiency, and is related to a proficiency test. The Norwegian construct, on the other hand, is intended for an achievement test, since it includes many specific skills and topical knowledge. In the analysis in chapter 6, these construct definitions will be actively used and referred to, and the method of the analysis is the subject of the following chapter.

5 Method

5.1 Chapter overview

The purpose of this thesis is to compare Vg3 level English examinations from Norway and Finland with regard to format and construct validity. How I went about doing the investigation will be explained in this chapter. I start by presenting and arguing for my choice of research design, also mentioning its weak and strong sides. This is followed by a brief description of how the Norwegian and Finnish samples were selected, and also how these two sets of tests differ from one another. It will then describe the method of content analysis, and how it has been applied in this case. The final subject of discussion is of validity issues.

5.2 Research design

The research design for this master thesis is based on the pilot study mentioned in the introduction (see section 1.3). I wanted to conduct a similar analysis, but on a larger scale, thus using roughly the same design. Since the idea is to look at differences between the examinations in Norway and Finland, this had to be a comparative study. And when documents such as these examinations are the objects of study, it is called document analysis or content analysis (Ary, Jacobs, & Sorensen, 2010). Because of the need to look closely at each test, and to carefully analyze each test task in relation to the construct, I decided to collect a small sample and conduct a qualitative analysis. In other words, I decided on a qualitative comparative study using content analysis as the most useful design for the present study. The perhaps main drawback of a qualitative design like this is that the results of the analysis are only valid for the specific tests in the sample. More elaboration on the validity of this method is included at the end of the chapter. In a qualitative approach, the sample is usually not selected randomly, but purposefully, and the next section will explain how the sample of examination papers for this thesis was selected.

5.3 Sample

The sample for this thesis was a selection of representative English examinations from Norway and Finland. My approach was to start by using criterion sampling. “In this type of

sampling, the researcher sets the criterion and includes all cases that meet that criterion” (Ary, Jacobs, & Sorensen, 2010, p. 431). One of the most important criteria in selecting the samples was comparability. Since the idea of the thesis is comparing tests from two countries, I needed to find examinations from those countries that were similar enough for comparison. This primarily meant that the tests should be at the same level, and serve the same purpose. Another criterion was that they be national examinations, since local tests may vary to a much greater degree. This narrowed down the search to the final examinations at the end of upper secondary school, since none of the examinations in Finland prior to this point is conducted at national level. As mentioned in the previous chapter, however, Norway offers two specialized English courses at this level, each of which have separate examinations. I selected the examinations from the course called “English literature and culture”, since the competence aims for this course corresponded best with the Finnish course descriptions, again using the criterion of comparability. Since the tests are given at the same level, and since both tests are examinations to provide grades needed for access to higher education, I would argue that they are comparable. There are, however, also some differences which will be mentioned in the section below.

At this point, I had to further narrow down the number of examinations to a number I could handle in a qualitative analysis. The Norwegian course of “English literature and culture” has only existed since 2009, and since there are two examinations every year, there were six tests in total to choose from when this project started in January 2011. However, and despite several correspondence attempts, I have not succeeded in finding the total number of the Finnish examinations of this kind. I can only be certain that these tests have been used since 2008, meaning at least eight in total (Yle, n.d.). Nevertheless, I systematically selected three of the most recent examinations from each country, over a time span of three years. Thus, from each of the years 2009, 2010 and 2011, I picked the spring semester version, resulting in a total of six tests. Further description of these tests is presented in chapter 6.

5.4 Content analysis

As defined above, construct validity is an overarching term for the justifiability of test use. Validation is gathering evidence to back up the justification arguments. There are many ways

to go about arguing for construct validity, and many ways to provide the evidence. Content analysis is one of them. As the theoretical chapter on assessment described, in order to justify the use of a test, every step from development to decision making must also be justified. In this paper, there will not be room to investigate every step of this process. Neither is it the point to create a comprehensive justification case for the use of these tests. The idea is rather to critically analyze some aspects related to their construct validity.

The validation method I have chosen for that use is called content analysis. Chapelle (1999) mentions this method as one of five different approaches to validity evidence. Content analysis is described as making “a logical analysis of the content to determine how well it covers the domain” (Ary, Jacobs, & Sorensen, 2010, p. 235). It is an analysis of content relevance; in other words, an analysis of the relationship between the test content and its construct definition. By comparing the test content and the test construct, it is possible to judge how well they correspond. Conducting such an analysis may provide “evidence for the hypothesized match between test items or tasks and the construct that the test is intended to measure” (Chapelle, 1999, p. 260). Since the test items are meant to measure the test construct, the hypothesis is that there is a match between the two. Doing content analysis is investigating whether, or to what degree, there is such a relationship.

Messick (1996) mentions two ways in which this relationship may be inconsistent. The terms he used to describe these two inconsistencies were *construct under-representation* and *construct irrelevance*. As figure 3 below illustrates, construct under-representation is when the construct includes some parts that are not adequately covered by the test. Conversely, construct irrelevance is when the test requires knowledge or skills that are not specified in the construct.

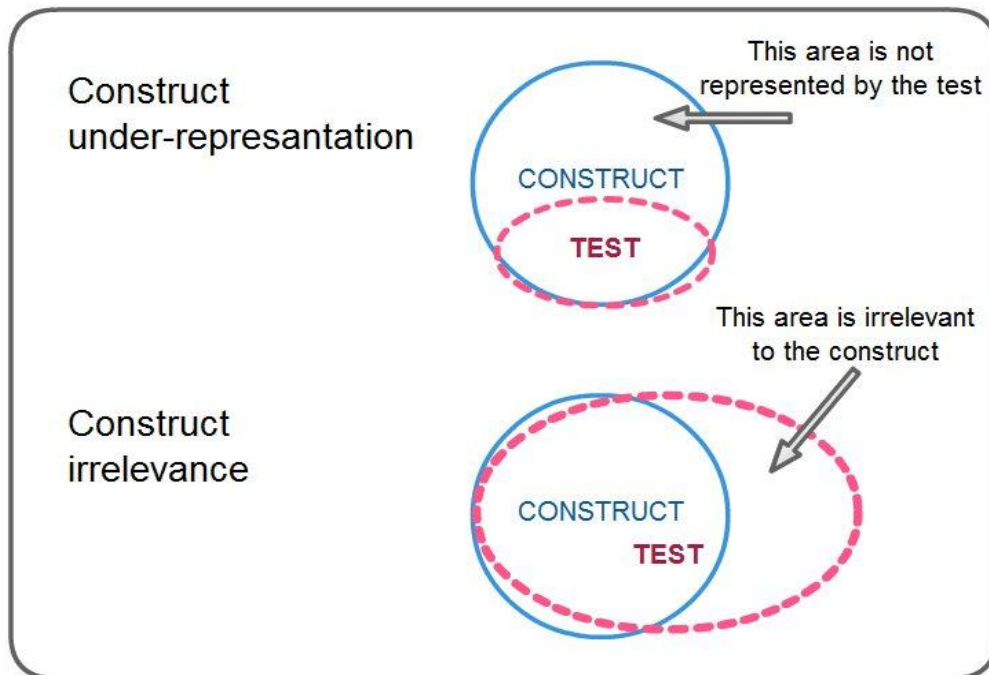


Figure 3: Illustration of construct under-representation and construct irrelevance

An example of the latter may be a listening comprehension test where a test taker is asked to write down his response to an oral text. If the test construct does not include anything about writing, this may cause a validity problem since writing is necessary in order to respond. A person who understands the oral text perfectly, but who cannot write in that language, may fail the test because of lack of a skill that was not supposed to be tested. This tells us that something is unreasonable. In order to resolve this validity issue, one would either have to add writing skills to the construct definition or change the way of responding to the task. Another way to improve the issue is to make it clear that the scores are by no means affected by the quality of the written responses. What is measured by the test is finally translated into a score. This means that what does not affect the scores is not really measured. This is an important point because it reminds us that the scoring of a test also needs to reflect the construct definition. This is also why the second part of my content analysis deals with the scoring procedures.

The other issue, construct under-representation, is when a test does not sufficiently cover all the major aspects of the construct. An example could be a language test needed in order to apply for a job in a telemarketing company. The test may only have included an oral interview where the examiner asks several questions, and where the score is based on the

grammatical correctness of the answers. This certainly covers parts of what is needed for professional phone conversations, but skills like taking initiative and keeping a conversation going is not included at all. The test is then not fully representative of the construct. It should be mentioned that it will always be the case that the construct is more comprehensive than the test. Therefore, a test cannot be expected to cover every detail of its construct. Hence, Messick (1996) puts it this way: “In the threat to validity known as construct under-representation (...), the assessment is deficient: the test is too narrow and fails to include important dimensions or facets of focal constructs” (p. 244). The key term here is “important dimensions”. If important dimensions of the construct are left out of the test or the scoring procedure, there is a validity issue of this kind.

Consequently, there are two areas which will be in focus in the following analysis. One area is the operationalization; the step from construct definition to the test tasks. The other area is the scoring procedures; its relation to the construct and its reliability. If there are inconsistencies between the construct definition and the assessment tasks, or the scoring procedure, it is crucial for the construct validity. Similarly, unreliable scoring will weaken the overall construct validity.

The procedure for this analysis will be as follows: The test tasks are firstly investigated to see what skills they require. For this step, the “accepted approach has been for expert raters to make judgments about the cognitive knowledge and processes they believed would be required for test performance” (Chapelle, 1999, p. 260). In this study, I will be the one making the judgments. Secondly, those required skills will be compared to the construct of the respective test. Similarly, the scoring procedures will be investigated and then compared to the relevant construct. An analysis of the scoring reliability will also be included in the last part.

As a tool for this analysis, I have developed a checklist of questions to ask for every examination or examination system. The checklist is my own, but is heavily based on checklist questions from Bachman and Palmer (1996), Messick (1996), Chapelle (1999) and

Johnson (2001). The questions are meant to provide the data needed for the procedure mentioned above.

OPERATIONALIZATION QUESTIONS

To what extent does the test task reflect the construct definition?

Analyzing test tasks:

- What is required for the test?

Comparing to construct:

- What is required for the test which is not included in the construct?
- What important dimensions of the construct are not required for the test?

SCORING QUESTIONS

To what extent do the scoring procedures reflect the construct definition?

Analyzing scoring criteria:

- What counts in the scoring procedure?

Comparing to construct:

- What counts in the scoring procedure which is not mentioned in the construct?
- What important dimensions of the construct are not taken into account in the scoring procedure?

Analyzing scoring reliability?

- Are the candidates allowed too much freedom in their responses?
- Are the tasks unequivocal, and the instructions clear and explicit?
- Is there agreement on what an acceptable response is?
- Are there multiple, independent raters?

The operationalization questions are individual for each examination, while the scoring questions concern the examination system as a whole. Thus, the first set of questions will be asked for every examination, and each task in every test will be considered. The last set of questions is only asked once for the Norwegian system, and once for the Finnish system, but obviously reflecting the reality of all of the tests in the sample.

5.5 Validity and limitations

As part of increasing the validity of this thesis, it is vital to address the issues that are, or can be a threat to its validity. One part of this concerns the sampling, and another part concerns the content analysis. “Sampling validity becomes an issue whenever a sample of texts differs from the population of the phenomena of interest” (Krippendorff 2012, p.336). In this case, as in most cases, the sample selection differs from the population, which means there is a chance

of sampling error. This study is obviously not a quantitative study, and, as mentioned before, it does not attempt to make statistical claims about a large population. However, it is desirable to be able to generalize the results to say something about the Norwegian and Finnish systems as such. Using a small sample size like this does not make it uncomplicated, but Krippendorff (2012) emphasizes the need to take into account the diversity within the population. As I have browsed through the tests that are not included in the sample, it is obvious that the population diversity is minimal. Each test strictly follows the pattern of the others, and differences are minor. Less diversity in the population means greater degree of generalizability. Krippendorff (2012) also mentions the proportion of the sample to the population. For the Norwegian tests, this factor is significant since fifty per cent of the population is actually included in the sample. This does not mean that the results of this analysis can be generalized to all the examinations in the population, and such strong claims will not be made. But I would argue that these facts add some validity to the sampling procedure, and opens for reasonably valid comments and discussion around the national examination systems as such, based on the results from this analysis.

The validity of this study is also related to the procedure of the content analysis. Content analysis is studying documents and is therefore “unobtrusive” and “nonreactive” (Marshall & Rossman, 1999, p.117). This means that it can be done without disturbing the setting of the study, and without changing the content of the studied objects. These are considered strong sides of this method, and are no threat to the validity. As with any method, however, there are some weaknesses that are important to be aware of. The moment a researcher doing content analysis decides to not only read the documents as they are, but also to infer meaning from them, there is a potential gap between the interpretation and the actual meaning of the text. This is a potential weakness of many types of qualitative designs, and means that one has to be careful to clearly communicate the reasons and rationale behind the interpretations made (Marshall & Rossman, 1999). Simply put, my interpretation may not be the same as that of any other researcher. Reviewing documents as science is therefore “highly dependent on the ability of the researcher to be resourceful, systematic, and honest” (Marshall & Rossman, 1999, p.135). If it is done in a systematic and honest way it will also be replicable and thus open to scientific criticism. In my case, in order to compensate for this weakness, I have

therefore been careful to clearly define my terms, attempted to be transparent in my method and follow a rigid structure in the analysis.

In discussing content analysis specifically of tests, it is said that “information about test content by itself is not a sufficient basis for score interpretations because it does not consider actual performance on tests” (Bachman, Davidson, & Milanovic, 1996, p.126). It is a weakness of this thesis that real test performance is not considered. This means that the results are largely based on my personal judgments:

This kind of analysis usually involves judgmental analysis of the skills and the processes required by test tasks. Although experts may experience difficulty in judging what an item measures, their judgmental analyses support the generation of hypotheses that can subsequently be tested by experimental or introspective studies (Grotjahn, 1986; Alderson 1990 as cited in XI, 2008, p.187).

As this quote underlines, making this type of judgment is challenging. That is also why it represents a weakness in the method. But the difficulty in making such judgments does not make the analysis completely useless. The results may still support hypotheses which can be further confirmed or unconfirmed by other studies. To sum up the essence, a final comment by Krippendorff (2012) on content analysis may serve well as a conclusion of this chapter:

The reason content analysts rely on face validity perhaps more than do researchers who use other methods of inquiry is that content analysis is fundamentally concerned with readings of texts, with what symbols mean, and with how images are seen, and of which are largely rooted in common sense, in the shared culture in which such interpretations are made, which is difficult to measure but often highly reliable at a particular time (Krippendorff 2012, p.330).

I have argued above that the use of content analysis, the checklist used, and the sample selection help provide useful and reasonably valid information. However, when all comes to all, it relies a whole lot on the common sense interpretations of language as done by the researcher. In that sense, it is up the researcher to be thorough and transparent in his procedures, and for the readers to judge the soundness of the arguments, which are to follow in the next chapter.

6 Results of the content analysis

6.1 Overview of chapter

This chapter presents the results of the content analysis described above. The findings are presented in sections 6.4, 6.5 and 6.6. The first one of these sections presents what I have found to be measured by the examinations. It shows what different skills are required to accomplish the test tasks, and what skills are rewarded by the scoring criteria. The second part compares the results of part one to the construct definitions outlined above in section 4.4. It presents the match or mismatch between what the construct says should be tested, and what I have found is actually tested. It will present the issue from two angles. The first is called construct irrelevance, which involves whether the tests measure something outside the construct. The other is construct under-representation, which is whether any important aspects of the construct are left out of the tests. In the third and final part, there is a presentation of the issues regarding scoring reliability. But prior to all of this, there will be a description of the examination formats and a description of the assessment documents from each nation.

6.2 Description of the examination formats

For each of the two examination formats, I will comment on the regulations, the instructions, other test input and the expected task responses. The most important differences between them will be summed up at the end. This will only be a general description of the patterns and common features of these tests, as details and examples from each one will appear below. Every test in the sample can also be found in its entirety in the appendices.

6.2.1 The Finnish English examinations

There are three Finnish English examinations in the sample, and, for convenience, I will refer to them as F09, F10 and F11, depending on the years they were used. The oldest one, F09, is available in Appendix 1, and was conducted March 13th 2009. The examination in Appendix 2, F10, was used on March 17th 2010, and the most recent examination in the sample, F11, was used on March 18th of 2011, and can be found in Appendix 3.

The time allotment for the Finnish English examinations is six hours, and without any particular required order in which to do the tasks. No notes or any help are allowed on the exam, thus reference to sources is not an issue. There is, however, some input for interpretation provided, but not related to any of the written production tasks.

These tests are divided into three main parts, each focusing on one or two particular language components. The first main part tests reading comprehension. It is a large part of the whole exam, but it will not be included in the analysis because this thesis concentrates on the productive skill of writing, and components relevant to that. It should be mentioned, however, that reading comprehension is tested in two tasks. The first one is a three-choice multiple-choice part with 25 items. The questions are detailed questions from three authentic and advanced English texts. The second one is five prompts with questions to a text where the response is given in the source language (Finnish/Swedish). However, as explained above, this first part will not be investigated further.

Part two is related to writing since it covers grammar and vocabulary. Grammar and vocabulary are tested in two different kinds of fill-in-the-blank tasks. The first has from 25 to 30 items to which a selected response is expected. The test taker can choose between four alternatives for each blank field in a coherent English text, and only one of the alternatives is correct. The second set of tasks includes between six and seventeen items, and it is a slightly different exercise. For each blank field there is usually a keyword provided. This word may be in the source language (Finnish/Swedish) and should then be translated and fitted to the right context. Alternatively, it could be a target language word in a basic form that needs to be conjugated correctly. In a few cases, there are no keywords at all, and the test taker is required to simply understand what is missing. The most recent exam, however, included a translation task instead of the kind of task just mentioned. It comprised an English text in which some sentences are in Swedish or Finnish instead, and where the task is to translate these sentences into appropriate English that fits in with the surrounding text.

The last part of the test papers is called production, and comprises composing a text of 150 - 250 words on one of four fairly general topics. Sometimes a setting or a genre is specified, but sometimes they only ask for a text addressing some general issue.

6.2.2 The Norwegian English examinations

There are also three Norwegian tests in the sample, which, in the following, will be referred to as N09, N10 and N11. N09 was conducted May 22nd 2009, N10 was conducted May 27th 2010, and the last one, N11, was used May 31 2011. They can be found in Appendices 4, 5 and 6 respectively.

These examinations have a maximum duration of five hours with no particular time sequence in which the tasks need to be done. Five hours is given to complete everything, and except for the internet and communication tools, almost all means of help or sources may be brought to the test. In other words, while books and notes are allowed on the test, all sources must be referred to when used in a response. Some source material is also appended to the test, and must often be used actively in order to carry out the tasks properly. All instructions and other input for interpretation are in written English, with the exception of non-language input, like the use of pictures in N10.

The main pattern in all the Norwegian tests is a structure of three main tasks. All of these tasks are similar in the sense that an extended written response is expected. This means that there are only writing tasks, and no multiple-choice or other limited response tasks. The difference between the three tasks concerns the length of the expected responses. Task 1 and 2 usually ask for a short response of only a couple of paragraphs. Task 1 is always subdivided into 1a and 1b, in which there are two quite specific questions about the features of a text. Task 2 is only subdivided in N11, and in that case, one only has to choose one of them. Task 2 often asks a more open question than in task 1. Finally, there is task 3 which presents the candidate with four topics about which he or she is to write a longer text. The options may vary in genre and content, but will always relate to some source text, some input for interpretation. The information required to write the essay may sometimes be provided in the

instruction and the source material, but this is not always the case. Some knowledge about culture and literature from the English-speaking world may also be expected.

6.2.3 Comparison of the two examination formats

The most noticeable difference between the two formats is the type of expected responses. A large part of the Finnish examination consists of multiple-choice or fill-in-the-blank tasks, while the extended writing is limited to texts of about 200 words. This stands in stark contrast to the Norwegian format, where the only response is written production, and a considerably greater amount of writing is expected. This means that the Finnish system prefers to test each language component separately, while the Norwegian one tests all of them together in the written text production. The other main difference relates to the topics of the written compositions. The Norwegian format expects the use of cultural and literary knowledge in some of the essay questions, while it is mostly general everyday topics that are dealt with in the Finnish tasks. These two points also suggest that there are different types of tests in each country. The Norwegian one can be categorized as an achievement test. Such tests are usually detailed, focused on objectives of the program and conducted at the end of a course (Brown, 1996). Moreover, the interpretation of scores is criterion-based, which also is more common for achievement tests (Brown, 1996). The Finnish examinations do not quite fit that definition. The scoring is norm-based and the essay topics are quite general. Although it has connection to objectives of a course program, and it is conducted towards the end of a school period, it could perhaps better be described as a proficiency test (Brown, 1996).

6.3 Description of the assessment documents

6.3.1 The Finnish assessment documents

The document containing the Finnish scoring regulations for these tests is called “Språkproven Föreskrifter och anvisningar” (Studentexamensnämnden, 2007). It is a general document concerning all language examinations in Finnish Upper Secondary School, but one of the chapters describes the assessment and scoring criteria for written language tests of this kind.

Before looking into how the written composition is assessed, I will briefly explain the much less complicated assessment of the Finnish grammar tasks. The outcome of the Finnish tests translates into a score of maximum 209 points, which gives the basis for the final grade. The grammar tasks count for only 40 of these points, compared to 99 for the production part. With regard to the grammar and vocabulary tasks, there are, as a rule, only two ways to score them, in which a correct answer gives one point, and incorrect answers give none. There is no gradation in any of these tasks except for the translation task in the most recent test, F11. For that task, one may be given from zero to three points depending on how close to an idiomatic and correct translation the response is.

Contrary to the grammar tasks, the assessment of the written production requires more specified criteria. In the table below, the criteria for top and bottom level performance from the Finnish assessment document are displayed.

Table 3: Scoring criteria for written production in Finnish English examinations

Points	Communicative ability	Content and structure	Linguistic range and accuracy
99, 97, 95, 92, 90	The writer is able to convey the message very clearly, almost completely idiomatically, fluently and nuanced. It is really easy to read the text.	The writer treats the subject in a very versatile, personal and consistent way, skillfully using cohesive devices of many kinds.	The writer has a very broad, versatile and idiomatic use of expressions, well suited for the situation. The writer masters these expressions really well.
(...)	(...)	(...)	(...)
15, 10, 5, 0	The writer lacks ability to convey the message.	The writer treats the subject completely inadequately.	The writer has a very basic use of expressions and even this is almost entirely incorrect.

Note. The content is from Studentexamsnämnden (2007), my translation

Table 3 is divided into three columns representing different aspects of the text to consider in the assessment. One aspect is communicative ability, which is defined as being able to convey a message. According to the descriptions in table 3, communication should be clear as well as

idiomatic and nuanced. Another aspect is content and structure. It deals with how the text treats the topic subject, and that the content is rewarded for being diverse and consistent. Linguistic range and accuracy is the last aspect, which describes the criteria for grammar and vocabulary. Using language expressions correctly and adapting them to the situation are the main criteria. Table 3 does not say anything about the relative importance of the three categories, but elsewhere, the assessment document affirms the following:

In the assessment of the performances, it is the ability of the examinee to present a message according to the task description, which weighs the most, i.e. communicative ability. Other criteria, such as content and structure, and linguistic range and accuracy, are primarily meant to support the holistic assessment done on the basis of communicative ability (Studentexamensnämnden, 2007, p.29, my translation).

Communicative ability is clearly the main emphasis, and being able to follow the task instructions also matters. The regulations also specify that deviating from the instructions is penalized by withdrawing points from the original score. If the composition changes completely direction or the task is fully misunderstood, points are lost according to the gravity of the mistake. If it is complete plagiarism or it treats a completely different subject than what is given, one will not receive any points from this task. Moreover, the Finnish instructions clearly state that the text should comprise 150 - 250 words. Deviation from this length also has consequences. If the text is more than ten per cent below the limit, five points are lost for each five percent missing. If the text is more than twenty five per cent too long, five points are lost for each twenty five percent extra.

6.3.2 The Norwegian assessment documents

In the Norwegian system, there is one general assessment document, as well as assessment guidelines that are specific for each test. The general document is called Assessment Guidelines (Nor.: “Vurderingsveiledning”), and the examination specific ones are called Rater Guidelines (Nor.: “Sensorveiledning”). The general guidelines refer to the grounds and basis for assessing the competence displayed in the tests. They state that the main basis for the assessment is the competence aims from the subject syllabus, that is, the test construct. The basic skills are also mentioned as an integrated part of the competence aims, and may therefore also be indirectly tested in the national examinations. Further, the general guidelines also make it clear that details are secondary to the overall impression in the assessment:

It must be emphasized that, in the assessment of the pupils' responses, details are to be seen in light of the totality. The final grade should be made on the basis of a total assessment of the entire paper (Udir, 2011b, p.4, my translation).

At the end of the Assessment Guide, there is a table outlining the general scoring criteria of a written text. I have included the description of the highest levels to illustrate:

Table 4: Scoring criteria for top grades (5-6) in the Norwegian English examinations

Level	Content	Written communication	Language
5 - 6	<p>The pupil gives precise, extensive and relevant answers to the tasks in accordance with the task instruction.</p> <p>The response shows signs of originality.</p> <p>If sources are used, they should be used properly, purposefully and with references.</p> <p>The pupil shows good insight and overview of the subject.</p> <p>The pupil discusses the relationship between form, content and style if relevant.</p> <p>The pupil analyzes linguistic features in different text types if relevant.</p>	<p>The pupil writes texts with good internal coherence.</p> <p>The pupil writes texts with logical paragraph division and good paragraph structure, using a theme sentence and supplementary sentences.</p> <p>Long texts have appropriate introductions and conclusions.</p> <p>The pupil writes texts of different genres.</p>	<p>The vocabulary is varied, nuanced and suited for the situation.</p> <p>The language of the pupil has good fluency and varied syntax.</p> <p>The pupil masters relevant terminology for analysis of literary texts, movies and other cultural expressions.</p> <p>The language of the pupil has no language mistakes disturbing the communication.</p> <p>The pupil uses cohesive devices in a versatile and appropriate way.</p>

Note. The content is from Udir (2011b), my translation.

Content, communication and language are the three aspects involved in table 4. At the highest levels of performance, as shown, the content must be relevant and precise, showing both overview and insight. It is also a criterion that the communication is structured and coherent and according to genre. Finally, the guidelines mention the linguistic aspect of writing,

describing nuanced and appropriate vocabulary, fluency and relevant terminology as the main criteria.

More detailed descriptions of how each task in every test is to be responded are found in the specified Rater Guidelines. They provide descriptions for every task, and how a response would look like at a high, medium and low level of performance. If relevant, the document also provides certain content that a proper response should contain. Although a typical response is described, the raters are also reminded to be open to different types of solutions as well.

6.3.3 Comparison of assessment documents

A common feature between the two assessment guidelines is the three aspects they operate with when assessing a written text. Despite some differences in the criteria descriptions, they both divide the criteria into the areas of communication, content and language. Furthermore, a comparison of the tables in the two assessment documents shows many parallels in each area.

First, the columns concerning language issues list many of the same aspects. The Finnish description uses fewer words, but the idea is the importance of broad vocabulary, correct and idiomatic language, and language that adapts to the situation. The Norwegian description covers the same aspects and more. It also mentions “relevant terminology for analysis of literary texts” which is unique for the Norwegian assessment. Second, in the content column, there are also some features concerning literature analysis as well as comments on source usage that only appear in the Norwegian document. The focus on how the content should be expressed also differs to some degree. While the Norwegian guidelines emphasize “precise, extensive and relevant” content, the Finnish one rewards “versatile, personal and consistent” content. This reflects the different focus on topical knowledge. Although the Finnish assessment document also considers the content in the scoring procedure, it is mostly a matter of whether the content is internally consistent and versatile. The Norwegian guidelines also consider whether the information is precise and relevant. Internal coherence is also mentioned in the Norwegian table, but in the communication category. Except for that, the two columns concerning communication cover mostly the same issues. In other words, the Norwegian table defines what it means to write clearly more explicitly, and it also rewards the ability to write in different genres.

A final point to mention where the two guidelines differ regards regulations on length and on following the instruction. The Norwegian documents simply state that extensive answers are rewarded. Contrary to the Finnish regulations, however, there is no rigid system for how length of response affects the score. Similarly, following task instructions is a criterion in both guidelines, but only the Finnish guideline has a system for how deviation from instruction affects the score. All in all, despite the differences mentioned, communicating the message, relating to the subject and correctly using linguistic expressions are the main criteria in both systems.

6.4 Test requirements: What do the tests measure?

A thorough investigation of the test requirements is essential for the following sections on construct irrelevance and under-representation. This section will illustrate what the tests measure by looking at the cognitive skills required for accomplishing the test tasks. In addition to this, aspects of the scoring criteria will be taken into account, since they also indicate what the tests measure. Altogether, the idea is to answer these questions from the checklist presented in section 5.4:

- What is required for the test?
- What counts in the scoring procedure?

6.4.1 Using the cognitive model

In order to analyze the tasks thoroughly, I will break the writing process into components that can be analyzed one by one. The cognitive model of writing from section 3.2 has been applied for that purpose. The model describes the use of external and internal input, which are processed in our working memory to produce the written output. External input is all the input from the examination papers. Using this input requires reading comprehension. According to the model, there are two internal inputs. Language knowledge is one, knowing how to write, and topical knowledge is the other, knowing what to write about. And finally, a certain cognitive level is required to turn these input sources into a relevant text. The following subsection will therefore present the requirements of the two test formats regarding the

following categories: reading comprehension, topical knowledge, language knowledge and cognitive level. A comparison of the Norwegian and Finnish results in each of these categories will be summed up together at the end of the section.

6.4.2 Reading comprehension

Every task, in both examination systems, contains an instruction in written English. This means that every task requires some extent of reading comprehension. However, this section will only present the required reading material besides the task instructions.

The Finnish examinations

The table below displays how much reading material is related to the Finnish tasks.

Table 5: Reading requirements in the Finnish English examinations

	F09	F10	F11
Task 2: Grammar & vocabulary	Non-fiction extract Article	Three articles	Non-fiction extract Parts of letter in English Parts of letter in L1.
Task 3: Production	-	-	-

As table 5 indicates, in task 3, the instruction is all the reading input there is. The grammar and vocabulary tasks, however, are tightly connected to the reading of authentic English texts. These texts are mainly articles with a formal and fairly advanced language. The kind of reading necessary to solve the tasks, however, depends on the task. In most cases, the task consists of finding a missing word in a sentence, either from among four alternatives, or based on a small clue. In order to do this correctly, one has to understand the text surrounding the blank field, but how much understanding is required? The examples below will demonstrate similar items requiring different levels of reading comprehension. Sometimes, only a quick look at the options and a look at the words next to the blank field is enough:

Above all they were responsible for / of / to / with the large sums of money...
(Appendix 3, p.14).⁵

⁵ In this and the next section (6.5), I have sometimes deviated from the APA 6 standard and also indented shorter quotes for the sake of legibility.

“Responsible for” is the most common combination, and a short look at the preceding verb phrase and the following noun phrase will quickly exclude the other prepositions. Other times, a close reading of the whole sentence is needed, but nothing more:

Anybody / Nobody / Everybody / All at Supreme Headquarters expected perfect conditions on D Day, least of all Eisenhower. (Appendix 1, p.14).

“Nobody” is the correct answer, but “everybody” could also seem like a possible option until the final tag phrase: “least of all Eisenhower” is read and understood. However, there are also some items that require an understanding of the larger context, such as these examples:

One of them told me the following day that his / her / their / a 18-year-old brother is serving in Iraq. (Appendix 2, p.12).

But enough with that / on top of everything / that was not all / notwithstanding. (Appendix 1, p.13).

In both these cases, several options could work if the sentences were decontextualized, as they are here. In total, eight out of eighty-five multiple-choice items, or approximately ten percent, require contextual information above sentence level. There is also a translation task in one of the tests, where being able to read simple sentences in English is useful, though not strictly speaking required.

Altogether the Finnish grammar tasks seem to require the ability to read and understand fairly advanced and formal English texts. Being able to understand details, as well as the ability to follow the context and understand the general meaning, are both necessary.

The Norwegian examinations

In comparison with the Finnish tests, quite extensive reading is necessary to accomplish the Norwegian ones. Only two out of twenty-two subtasks do not require reading besides the instruction. Except task 1b and task 2 in N10, all other tasks require reading of a novel excerpt, an article, a poem or another text of some kind. An overview of this is presented in table 6.

Table 6: Reading requirements in the Norwegian English examinations

	N09	N10	N11
Task 1	Article	Academic text one paragraph	One paragraph
Task 2	Poem	(Picture)	One paragraph / poem
Task 3a	Novel excerpt	Poem	Novel excerpt
Task 3b	Novel excerpt	Lyrics	Novel excerpt
Task 3c	Poem	Novel excerpt	Lyrics
Task 3d	Novel excerpt	Two citations	Novel opening

The texts used in the Norwegian tests vary from formal and fairly advanced literature to texts that are simpler, using quite colloquial language. For the most part, the texts are fiction or poetry, and sometimes, poetic imagery and cultural allusions are used in the poems. A passage from one of those poems is displayed below:

They still are ranged along the roads
 plagued by legionnaires
 false windmills and demented roosters
 They are the same people
 only further from home
 on freeways fifty lanes wide
 on a concrete continent
 spaced with bland billboards
 illustrating imbecile illusions of happiness (Appendix5, p.9).

To understand such metaphorical expressions, a quite advanced level of reading comprehension is needed. In this, as in most of the Norwegian tasks, the general idea of the text must be understood as well as particular detailed information. In poetry analysis, one would have to comment on the theme and message of the poem, which requires a general understanding. But in order to comment on linguistic features, one has to read for details as well. This is typical for many of the Norwegian essay topics, as well as for the smaller tasks, as in this example:

Read the text and answer a) and b):

a) Summarise in your own words the author's message in one or two sentences.

b) Point out at least three different linguistic devices used by the author to argue his point, and explain how each of them enhances the message the author wants to convey (Appendix 6, p.6).

Pointing out certain linguistic features of a text, and commenting on their relation to the message or genre is the typical first task in all of the Norwegian tests. That is to say, not only is reading required in virtually all the tasks, the nature of most of the tasks require thorough analytical reading, and also sometimes of quite metaphorical language. On the other hand, one may argue that use of dictionaries may reduce the need for understanding of advanced vocabulary.

6.4.3 Topical knowledge

What do the candidates of these tests need to write about? What knowledge is necessary to respond successfully to these tasks? We have already mentioned that the Norwegian achievement test is more focused on content than the Finnish proficiency test, and that is also what this analysis reveals. Tasks requiring specialized topical knowledge are indicated in tables 7 and 8 below. The other tasks are labeled “general” if only common, general knowledge from the writer's memory is required. It may still be the case that other knowledge is required, but that sufficient information is provided in the test. It may also be the case that personal opinions or invented stories are the subject of the composition. Sometimes, it may be difficult to judge whether a topic should be considered general or specialized, in which case the topic is put in brackets.

The Finnish examinations

The Finnish grammar part does not really require anything but language knowledge. Going through each item, though, I have come over two cases where knowledge of the world does help. It is perhaps not possible to argue that this knowledge is the only way to differentiate between the options, but some basic knowledge of the photosynthesis would for instance be useful in this example:

[Bamboo] also demands / produces / needs / requires considerably more oxygen from carbon dioxide than trees (Appendix 2, p14).

Nonetheless, for the Finnish format, the question is rather whether topical knowledge is required for the written production tasks. Table 7 below affirms that it is not, showing the required topical knowledge for each of the four alternative topics in each test.

Table 7: Topical knowledge required in the Finnish composition tasks

	F09	F10	F11
Task 3.1	(Genes/environment)	General	General
Task 3.2	(Role-play)	General	General
Task 3.3	General	(A world changer)	General
Task 3.4	(Extreme sports)	James Bond	General

In the table, most of the tasks are labeled “general”. A typical example of one of these is writing a birthday speech to one of your grandparents (Appendix 1). This gives the writer complete freedom to fill in content of own choice. Knowledge about birthdays, parties and grandparents is considered very general. Another example is this task:

You live in England and there are plans to build a new motorway, which will pass very close to your home. Write a letter to be published in a local paper giving your opinion about the plan (Appendix 2, p.18).

When the whole task is to express your own opinion on a given issue, there cannot really be any requirements concerning content. There is only one obvious example of specialized topical knowledge in the Finnish sample:

James Bond – a hero of the past?

A film magazine has invited its readers to take part in a writing competition under the title above. The winning entry will be published. Write your article about James Bond for the magazine (Appendix 2, p.18).

Writing an article about James Bond for a film magazine is certainly difficult without some specific knowledge about the movie character. The rest of the Finnish topics are general, although some are discussable. Borderline topics include for instance role-plays and extreme sports. These are quite common subjects among young people, but not many would perhaps know all the details. In any case, although one would count all the borderline cases as specialized topics, table 7 clearly shows that there is always at least one topic where they can be avoided in every test. Consequently, it seems fair to conclude that although some knowledge may be helpful, strictly speaking no specialized knowledge is required in the Finnish format. This is supported by the focus of the scoring criteria, where the content is not rewarded for being factual, but only for being consistent and versatile.

The Norwegian examinations

Analyzing the need for topical knowledge in the Norwegian tests is slightly more complex, but table 8 below gives a simplified overview to emphasize the important findings.

Table 8: Topical knowledge required in the Norwegian examinations

	N09	N10	N11
Task 1	General	General	General
Task 2	Presidential election	General	General
Task 3a	<i>General</i>	General	General
Task 3b	<i>General</i>	General	(British society)
Task 3c	Romantic poetry	(Gender roles)	General
Task 3d	General	(Vampire literature)	General

First, we look at the two tasks marked in bold in table 8, which clearly require specialized topical knowledge:

Write two or three paragraphs in which you discuss how the poem (...) below expresses some of the feelings evoked by the election of Barack Obama as US President in 2008 (Appendix 4, p.6).

The second task asks the test taker to:

Write an analysis of the poem below pointing out its typical Romantic features (Appendix 4, p.7).

The poems referred to are provided, but the test taker is expected to know some relevant information about the presidential election or Romantic literary features. This information is not typically defined as general or common knowledge; neither is it to be found in the poems or other source texts provided. The rater guidelines for these two tasks also specifically ask for this knowledge to be demonstrated by the examinees (Udir, 2011b). Additionally, table 8 includes some borderline cases marked in brackets, either because the topic is fairly general, or because it is only helpful, not required, to know about the topic.

Second, a comment is necessary concerning the tasks labeled “general”. Two of them are uncomplicated and are indicated in italics in table 8. They do certainly not require any

specialized topical knowledge. They are both short stories in which the content can be invented freely, and where the rest of the information needed is provided in the source texts. However, all the other tasks are some kind of textual analyses. In these tasks, the rater guidelines do not specify any topical knowledge to be shown, but they do require some knowledge of linguistic features, like genre features, literary devices and specialized terminology for literary analysis. An example would be task 3d in N11:

“Vedlegg 3” presents the opening of Meg Mullins’s novel *The Rug Merchant*.

Write an essay in which you discuss how effective the opening of the novel is, what we learn about the two characters Ushman Khan and Mrs. Roberts and what we learn about the relationship between them. Use examples from the extract in your essay (Appendix 6, p.8).

The main content of the response is information retrieved from the novel opening, which is provided. The examinee will not need to write about anything besides the novel excerpt, but knowledge of literary devices is needed in order to discuss the effectiveness of the opening. The rater guidelines, when describing a response at a high level, put it like this:

In a response at a high level, the pupil discusses the introduction of the novel with originality and focus. The pupil gives relevant examples from the text, and uses appropriate terminology. The pupil writes thoroughly about the characters and the relationship between them. The text is well structured. The language is nuanced and versatile (Udir, 2011c, my translation).

All the information needed for a good response seems to be found in the test itself. There is no mention of specific content that should be included. But in order to find relevant examples from the text, one has to know something about the characteristics of a novel. For the most part, it has to do with knowledge in order to analyze literature. In one of the obligatory tasks, though, an advertising poster is shown, and the instruction is as follows:

Write two or three paragraphs in which you interpret the message of the advertisement above, and discuss how the message is conveyed (Appendix 4, p.6).

In this case, it is not analysis of literature which is required, but the analysis of a picture. Certain knowledge of the features and terminology of pictures and art is therefore needed here.

All in all, it may seem like topical knowledge is not such a large part of these examinations after all. Disregarding task 2 in N09, it is possible to avoid tasks requiring anything but knowledge of linguistic features. The conclusion is that explicit knowledge of linguistic features is required, but other areas of topical knowledge about cultural and social issues are barely touched upon in two of the three tests. What is more, since books and notes can be brought to the test, one may not have to rely on having this knowledge at all, but only refer to valid sources. I will return to this issue in the section on construct under-representation below.

6.4.4 Language knowledge

Obviously, all these tasks require a certain amount of language knowledge. In order to communicate something about the topic, language knowledge is necessary. Language knowledge is defined as communicative competence, including linguistic, socio-linguistic and pragmatic competence. The linguistic aspect concerns grammar and textual cohesion. The socio-linguistic and pragmatic components involve adapting the language to the social context in order to convey the intended meaning. Some grammatical competence is needed in any written exercise, and textual competence is needed in the production of any coherent text. The linguistic component is therefore measured in both tests. When it comes to the socio-linguistic and pragmatic aspects, however, it is much more challenging to define whether it is required or not. For that reason, I have only focused on two relevant components of socio-linguistic competence: genre knowledge and situation appropriate language. Tables 9 and 10 below indicate the genre specifications of the tasks, and if a genre is marked in bold, it means that a communicative situation is also specified.

The Finnish examinations

One can hardly say that any socio-linguistic knowledge is necessary to choose the right option in the multiple-choice tasks. As we have seen, an understanding of the textual context may sometimes be required, but not of the social context. Those tasks only measure linguistic competence, mostly grammatical, and perhaps textual to some degree. Table 9 below gives an overview of genres for the rest of the tasks:

Table 9: Genre specifications in the Finnish English examinations

	F09	F10	F11
Task 2.2	-	-	personal letter
Task 3.1	none	speech	none
Task 3.2	none	letter to the editor	none
Task 3.3	speech	none	letter to the editor
Task 3.4	none	article	none

Table 9 makes it clear that the Finnish composition topics vary from year to year with respect to genre specifications. In 2010, all but one task were genre-specific, while the other two years, it was the other way around. Consequently, it is possible to avoid the genre-specific tasks in all these tests, but even if a genre task is chosen, genre knowledge is not measured. This is because the scoring criteria do not take genre features into account. There is no criterion requiring the written text to adhere to genre specifications. On the other hand, as the assessment guidelines affirmed, use of language suited to the situation is rewarded. And most of the Finnish tasks that provide genre specifications also provide a setting, a purpose and an intended audience:

You will be participating in European Youth Week and will give a speech on a topic you find important to all young Europeans. Write this speech (Appendix 2, p.18).

The majority of tasks, however, do not specify any context at all. How could a rater judge whether the language is well adapted to the situation in a task like this:

When you want to celebrate something with your friends, do you yourself prepare the food or do you prefer some other arrangement? Why? What does your “menu” consist of? (Appendix 3, p.18).

One may conclude that genre knowledge is not tested in the Finnish examinations, while situation appropriate language is to some extent. The only place to really demonstrate such competence, however, is the composition tasks which are relatively short, and sometimes without enough contextual information.

The Norwegian examinations

In the Norwegian examination format, the situation is quite different. The scoring criteria reward both that “the pupil writes texts of different genres”, and that the pupil uses vocabulary which is “suited for the situation” (Udir, 2011b, p.5). As is evident from table 10, most Norwegian tasks also include genre specifications, contrary to what was found in the Finnish tests.

Table 10: Genre specifications in the Norwegian English examinations

	N09	N10	N11
Task 1	none	literary critique	none
Task 2	none	none	(copy attitude) / none
Task 3a	character analysis	poetry analysis	essay
Task 3b	short story	short story	essay
Task 3c	poetry analysis	essay	essay/analysis
Task 3d	essay	essay	essay

The table reveals a general pattern where task 1 and 2 usually do not specify a genre, while every single alternative in task 3 does. As this example shows, using the genre features is sometimes even explicitly required in the task instruction:

Write a short story about the character in the lyrics “Misunderstood” (“Vedlegg 3”) that reveals what type of person he is or represents. In your short story, you should make use of the features of this genre (Appendix 5, p.7).

The need for genre knowledge seems to be quite obvious. There is, however, no explicit mention of situations or intended audience in these tasks, although situation appropriate language is a criterion in the assessment guidelines. The specified genres will usually require a certain register of appropriate language, but the communicative situation is still missing.

All in all, it seems like some of the socio-linguistic aspect is taken into account and measured in the Norwegian format, particularly the component of genre knowledge. Situation

appropriate language is also measured since it is included in the scoring criteria, but the situations are only implied by the genre specifications.

6.4.5 Cognitive level

Following the cognitive writing model mentioned above, we see that all the input elements have been discussed. The external input was reading comprehension, and input from internal memory was language and topical knowledge. The final component is the processing of this input into written production. What level of cognitive processing is required in these tasks? In chapter 3, section 1, this taxonomy of cognitive processing was introduced:

1. Reproduce
 2. Organize / reorganize
 3. Invent / generate
- (Vähäpääsi, 1982 in Weigle, 2002, p. 8)

The Finnish grammar tasks, where filling in a form is all it takes, are at the lowest level, reproduction. Tables 11 and 12 list up the types of writing asked for in the other tasks, giving an indication of the cognitive levels they demand. Judging what cognitive level is required for any given task may be challenging, but a quite distinctive tendency appeared in the analysis.

The Finnish examinations

I only found three different types of writing in the Finnish tests, indicated by table 11.

Table 11: Types of writing tasks in the Finnish examinations

Writing tasks	Frequency
Give opinion and explain	8
Give a speech	2
Discuss	2

Eight out of twelve Finnish tasks consist of giving a general explanation and a personal opinion as in this example:

Consuming is the trend of the day. What are real necessities? What could we do without? How could we save money? (Appendix 3, p.18).

Table 11 further shows that two of the tasks ask for a written speech. A speech in a formal setting may include a discussion, while an informal speech could be mainly commenting, narration and personal opinions, and one of each is represented in this sample. Two of the tasks require discussion, as in this example:

Is it necessary to have shops open all days of the week? What advantages or disadvantages has it brought? Write a letter to the editor of a newspaper (Appendix 3, p.18).

It is a letter to the editor, and one would expect a discussion of several sides of an issue. The cognitive demand seems higher since the writer has to relate to a particular genre, as well as arguing for and against Sunday shopping. Yet, as long as the content is versatile and consistent, the scoring criteria do not seem to consider the level of discussion in the text (Studentexamensnämnden, 2007). Nor is the ability to follow the specified genre considered in the scores, as mentioned above. The conclusion is that the cognitive requirement hardly surpasses level two in the taxonomy above, which is the organizational level.

The Norwegian examinations

Again, there is a contrast to what is found in the Norwegian tests. Not only does the Norwegian format allow for several different types of writing, but there is also a majority of high-level cognitive tasks, as displayed below in table 12.

Table 12: Types of writing tasks in the Norwegian English examinations

Writing tasks	Frequency
Discuss	8
Analyze	5
Compare	3
Narrate	2
Summarize	1
Comment	1

The two most frequent categories in this table are *discuss* and *analyze*. They count for almost twice the amount of the other tasks. Next is *compare*, *narrate*, *summarize* and *comment*. Summarizing is a typical example of reorganization, and together with narration and commenting, they probably fall under the organizational category. The most frequent tasks,

however, require a higher level of cognitive skills. When analyzing a poem, for instance, it is not only a matter of organizing known knowledge. One is presented with a new text, and has to gain knowledge from that text to use in the written production. Moreover, the scoring criteria expressly emphasize these qualities of a text:

- The pupil discusses the relationship between form, content and style if relevant.
- The pupil analyzes linguistic features in different text types if relevant.

(Udir, 2011b, p.5).

Consequently, it seems that the required cognitive level is relatively high, perhaps closer to level three in the taxonomy, “invent / generate”, also known as knowledge transformation.

6.4.6 Summing up so far: Comparison of requirements

The next section will see how well the results so far match the construct definitions. A summary of the two examinations systems and a comparison of what they measure will be useful before moving on.

Table 13: Comparison of what the Finnish and Norwegian English examinations measure

	Finnish examinations	Norwegian examinations
Reading compr.	<ul style="list-style-type: none"> General reading comprehension Reading for detail Understanding formal, advanced Language 	<ul style="list-style-type: none"> General reading comprehension Reading for detail Understanding formal, advanced language Understanding poetic imagery
Topical knowledge	<ul style="list-style-type: none"> Only general topical knowledge 	<ul style="list-style-type: none"> Specialized topical knowledge such as: <ul style="list-style-type: none"> - of the presidential election - of the Romantic literary period - of literary features - of picture interpretation
Language knowledge	<ul style="list-style-type: none"> Grammatical knowledge Textual knowledge Situation appropriate language 	<ul style="list-style-type: none"> Grammatical knowledge Textual knowledge Genre knowledge Situation appropriate language
Cognitive level	<ul style="list-style-type: none"> Explain and give opinion Give speech 	<ul style="list-style-type: none"> Discuss Analyze Compare Narrate Comment Summarize

The table above reveals that the Norwegian tests cover more than the Finnish ones in practically every area I have studied. This is true for reading comprehension, topical knowledge, language knowledge and cognitive level.

In both formats, I have found reading to be an essential part of the tests. Reading comprehension is not what they primarily want to measure, but they do it indirectly because it is required to properly solve the writing tasks. Although the ability to understand general meaning and details is measured in all the tests, it turns out that the Norwegian tests include reading to a greater extent, and also a greater variation of texts to read. At times, the Norwegian texts also demand more advanced reading skills than do any of the Finnish ones.

Second, I found that they differ when it comes to topical knowledge, but not as much as I had expected. The Finnish format succeeds in making writing tasks that only require general topic knowledge, and the Norwegian one includes some knowledge about culture and literature. It was shown, however, that most Norwegian tasks measure knowledge that borders on being language knowledge. The discussion chapter below will elaborate more on this distinction. When disregarding those tasks, the remaining topical knowledge turned out to be limited. In fact, two of three tests allowed the test taker to avoid topical knowledge altogether.

Third, there was a difference in what type of language knowledge the two examination systems measure. The essence of the finding was that genre knowledge is required to a much greater degree in the Norwegian format. The Finnish scoring criteria do not ask for genre features; neither do most of the composition tasks. In contrast, all the Norwegian composition tasks specify a genre, and using genre features is also a scoring criterion. When it comes to adopting the language to the situation, it is a part of the scoring criteria in both test formats. However, the Norwegian tasks do not specify a communicative context, only the expected genre. And only a few of the Finnish tasks provide a contextual situation. In other words, the socio-linguistic component is measured to some extent in the Norwegian tests, but hardly at all in the Finnish tests.

Finally, I also found the cognitive level measured to be highest in the Norwegian tests. The Finnish tasks are mainly limited to writing explanatory texts and sharing views and opinions. While these may be touched upon in some Norwegian tasks as well, the majority of them require a cognitive level of discussion and analysis.

6.5 The relationship between requirements and construct: Do they measure what they should?

6.5.1 Construct irrelevance

In this section we will look at the following questions from the checklist, first regarding the Finnish examinations and then the Norwegian examinations:

- What is required for the test which is not included in the construct?
- What counts in the scoring procedure which is not mentioned in the construct?

In other words, do the construct account for everything we have found to be measured by the tests, or is there any observable construct irrelevance?

The Finnish examinations

A quick look at table 12 below reveals particularly one questionable area in the Finnish tests regarding construct irrelevance. However, before this is explained, I will also demonstrate the construct adequacy in the other areas.

Table 14: Construct irrelevance for the Finnish English examinations

	Test requirements	Included in construct
Reading comprehension	General understanding	Yes
	Detailed reading	Yes
	Read non-fiction / articles	Yes
Topical knowledge	General topical knowledge	Yes
Language knowledge	Grammatical knowledge	Yes
	Textual knowledge	Yes
	<i>Situation appropriate language</i>	No
Cognitive ability	explain	Yes
	give opinion	Yes
	give speech	Yes

Reading comprehension is the first area in table 14, but the construct definition from chapter 4 does not mention this skill. This is because we have focused on writing, and left out the descriptions of reading comprehension although they are an integrated part of the total test construct. It is therefore fair to say that the construct includes the reading requirements as listed above. The reading proficiency scales states that a language user at this level:

- Can identify the meaning of a text and locate several different details in a long text
- Can read a few pages of text independently (newspaper articles, short stories, popular fiction and nonfiction, reports and detailed instructions)

(Finnish National Board of Education, 2004, p. 247).

Every aspect is included there; understanding details and meaning, as well as the reading of formal non-fiction literature.

Topical knowledge is hardly even an issue, since we have found that no specialized knowledge is really required. Even so, the construct clearly includes this aspect also speaking of writing “clear and detailed texts about a variety of areas of personal interest and about

familiar abstract topics” (Finnish National Board of Education, 2004, p. 247). This is just the type of topics we have found in the tests as well, causing no construct irrelevance.

The next area, language knowledge, is where the inconsistency is found. The grammatical and textual components of language knowledge are included in the construct in these formulations:

- Has a fairly good command of orthography, grammar and punctuation and errors do not lead to misunderstandings.
 - Can use broad vocabulary and demanding sentence structures together with linguistic means to produce a clear, cohesive text
- (Finnish National Board of Education, 2004, p. 247).

It is questionable, however, whether situation appropriate language is adequately defined in the construct. There is one criterion related to genre knowledge, but the aspect of adapting to the situation is not included:

- Can write (...) routine factual messages and more formal social messages (reviews, business letters, instructions, applications, summaries)
- (Finnish National Board of Education, 2004, p. 247).

Since the scoring criteria specifically mentions that good language ought to be “well suited for the situation” (Studentexamensnämnden, 2007, p.42), one would expect to find a corresponding description in the construct, but there is none.

Finally, there is a good match between the cognitive level required in the test and what is described in the construct. The most frequent type of writing in these tests is to *explain* and *give one's opinion*. The construct states that writing at this level includes being able to “express *information* and *views* effectively in writing”. It could not be more obvious that what the tests ask for in this respect is also what the construct requires.

The Norwegian examinations

I have not found any severe inconsistencies regarding construct irrelevance for the Norwegian tests, but some question marks have appeared. They will be explained below as I go through each of the four categories in the leftmost column in table 15.

Table 15: Construct irrelevance for the Norwegian English examinations

	Test requirements	Included in construct
Reading comprehension	General understanding	Implicitly + as basic skill
	Reading for details	Implicitly + as basic skill
	Read non-fiction/articles	Implicitly + as basic skill
	Poetic imagery	Implicitly
Topical knowledge	Literary knowledge	Yes.
	Romantic literary period	In general terms
	Presidential election	Yes
	Picture interpretation	In general terms
Language knowledge	Grammatical knowledge	Mostly implicitly
	Textual knowledge	Mostly implicitly
	Genre knowledge	Yes
	Situation appropriate language	Yes
Cognitive ability	Argue / Discuss	Yes
	Analyze	Yes
	Compare	Implicitly
	Narrate	In general terms
	Summarize	Yes
	Comment	Yes

As table 15 shows, practically all test requirements are included in the construct, although some are less explicitly stated than others. Reading comprehension is, for instance, not explicitly mentioned in the Norwegian construct at all. The competence aims deal mainly with productive skills. To “use a nuanced, well-developed and precise vocabulary” is one of the

aims, but it does not speak of understanding advanced vocabulary. However, the following parts of the construct certainly involve reading comprehension:

- interpret literary texts (...) from a cultural-historical and social perspective
- interpret a (...) selection of texts (...)
- analyse at least two lengthy works of fiction

(Udir, 2006, p.6).

Such interpretation and analysis presuppose a basic reading skill, and reading is also defined in the syllabus as one of four basic skills. The basic skill of reading at this level is defined as follows:

Being able to read in English involves understanding, exploring and pondering demanding texts, thereby gaining insight across cultures and special fields. This is an integral part of practical language skills. It also involves the ability to choose a reading strategy suited to the intended purpose (Udir, 2006, p.3).

We see that it covers also demanding texts as well as cultural and special fields, which arguably includes reading of advanced language and even poetry. Since it involves reading strategies, it also covers both general reading and reading for details.

Some of the question marks appear in the analysis of topical knowledge. I have identified four elements of required topical knowledge, half of which are clearly described in the construct.

Knowledge of literary features is mentioned several times, like in this competence aim:

“elaborate on and discuss distinctive linguistic features of texts from different genres, from different periods and regions” (Udir, 2006, p.6). To expect knowledge about the presidential election is justified since the construct includes the skill to “elaborate on and discuss current issues in international culture and the news media” (Udir, 2006, p.6).

Topical knowledge of the “Romantic literary period” and “Picture interpretation” is only included in the construct in general terms, as table 15 indicates. In other words, the Romantic period is not mentioned anywhere specifically, only included in a general competence aim like this:

- interpret a representative selection of texts from literary-historical periods in English literature, from the Renaissance up to the present time (Udir, 2006, p.6).

Similarly, knowledge about picture interpretation is only mentioned if it is included as “other forms of expression” in these competence aims:

- have a command of the terminology needed for analysing works of fiction, films and other aesthetic forms of expression
 - analyse and assess a film and a selection of other artistic forms of expression within English-language culture
- (Udir, 2006, p.6).

Regarding language knowledge, there is not much to mention. It is not always explicit in the Norwegian construct, but grammatical, textual and socio-linguistic features are all covered. These three competence aims contain the language related elements of the construct:

- use suitable language, appropriate to the situation, in oral and written genres
 - use a nuanced, well-developed and precise vocabulary to communicate on literature and culture
 - produce texts in a variety of genres with clear content, appropriate style, good structure, and usage that is precise and accurate
- (Udir, 2006, p.6).

Both genre knowledge and situation appropriate language are mentioned. So are vocabulary, accurate language and good structure.

Finally, a comment is needed on the cognitive level. The types of writing listed in the table above correspond well with the construct definition. Except for narration and comparison, the construct refers to all of the types explicitly. Narration is included in a general way since knowing how to write different genres is one of the defined aims. Comparison is implied if the pupil is able to do the following:

- summarize, comment on and discuss *differing viewpoints* in fictional texts
- (Udir, 2006, p.6).

Construct irrelevance does not seem to be an immense problem in any of the examination formats. Most of what was found to be measured by these tests is also described in the test construct, implicitly or explicitly. In the Finnish examinations the construct does not mention situation appropriate language although it is a scoring criterion, while for the Norwegian examinations, one may speak of a potential issue regarding topical knowledge which is only generally stated in the construct.

6.5.2 Construct under-representation

Just as damaging to the overall construct validity of a test as construct irrelevance is construct under-representation. The checklist questions relevant to this section are as follows:

- What important dimensions of the construct are not required for the test?
- What important dimensions of the construct are not taken into account in the scoring procedure?

We recall that the construct will usually contain a broader scope than a given test, but that if an important dimension of the construct is left unmeasured, the construct is under-represented. Here follows an analysis of the constructs, and how the tasks and the scoring criteria measure the different construct elements.

The Finnish examinations

The components of the Finnish construct definition are listed in table 16 below. The descriptions remind us that this concerns a proficiency test since most of them are general skills used in any writing task. The right column indicates how many tasks in the Finnish sample that measure the different construct components, and the total number of composition tasks is twelve.

Table 16: Number of tasks where each part of the construct is measured in the Finnish English examinations

Construct definition	Number of tasks where this is measured
Can write clear and detailed texts about a variety of areas of personal interest and about familiar abstract topics	12
Can write routine factual messages and more formal social messages (reviews, business letters, instructions, applications, summaries).	5
Can express information and views effectively in writing and comment on those of others.	12
Can combine or summarise information from different sources in his/her own texts.	0
Can use broad vocabulary and demanding sentence structures together with linguistic means to produce a clear, cohesive text.	12
Has a fairly good command of orthography, grammar and punctuation and errors do not lead to misunderstandings.	12++

It is apparent from table 14 that most of the construct elements are represented in all the twelve production tasks. The last element in the table, concerning orthography and grammar, is also covered in the selected response tasks. Further, it shows that only five tasks measure genre knowledge, although the construct states that certain knowledge of basic formal and social genres is expected at this level. Even more striking, however, is the absence of tasks requiring a summary of different texts. None of the writing tasks refers to any sources at all, and none of them involves summarizing source information. I also examined the reading comprehension tasks (although they are otherwise excluded from this study) to see if any of them required summary of different sources, but found that they did not.

The Norwegian examinations

The Norwegian construct, as outlined in chapter 4, consists of three main areas. In order to look at construct under-representation in the Norwegian tests, it is useful to consider one such area at the time, to see in how many tasks each component is measured. In these tests, the total number of tasks is eighteen.

Table 17 presents the results for the area called “language and language learning”. This area describes the knowledge about linguistic features and terminology which was mentioned in the previous section about topical knowledge.

Table 17: Number of tasks where language and language learning aims are measured in the Norwegian English examinations

Main area: Language and language learning	Number of tasks where this is measured
elaborate on and discuss the relationship between form, content and stylistic register in sentences and text	9
have a command of the terminology needed for analysing works of fiction, films and other aesthetic forms of expression	7
elaborate on and discuss distinctive linguistic features of texts from different genres, from different periods and regions	8

Discussing form and content, knowing the terminology for analysis and discussing genre features are the three elements in this table. As table 17 shows, each of them is measured in a reasonably high number of tasks, so nothing appears to cause construct under-representation in this area.

The next area is communication, presented in table 18:

Table 18: Number of tasks where communication aims are measured in the Norwegian English examinations

Main area: communication	Number of tasks where this is measured
use a nuanced, well-developed and precise vocabulary to communicate on literature and culture	15
elaborate on and discuss lengthy and linguistically demanding discourses with general, specialized and literary content	12
use suitable language, appropriate to the situation, in oral and written genres	14
summarize, comment on and discuss differing viewpoints in fictional texts	3
produce texts in a variety of genres with clear content, appropriate style, good structure, and usage that is precise and accurate	13

From this table, it is noticeable that one element has a considerably lower number than the rest. Discussing viewpoints in fictional texts is clearly not the most common task. Otherwise, table 18 indicates that the communication aims are generally measured to a great degree in all the tests. Using a nuanced vocabulary, discussing demanding discourse and writing in different genres are all well represented. One point is questionable, however, regarding use of situation appropriate language. According to the construct, it should be measured, and the table indicates fourteen tasks that measure it. However, thirteen of those fourteen tasks are counted because they specify a genre, but not specifically a situation, or social context. The same thirteen tasks therefore also appear in the component of genre knowledge. This issue will be discussed in the next chapter.

As for the main area called “culture, society and literature”, there are surprisingly few tasks designed to measure this aspect, clearly indicated by table 19 below.

Table 19: Number of tasks where culture, society and literature aims are measured in the Norwegian English examinations

Main area: Culture, society and literature	Number of tasks where this is measured
interpret a representative selection of texts from literary-historical periods in English literature, from the Renaissance up to the present time	1
analyse at least two lengthy works of fiction	3
analyse and assess a film and a selection of other artistic forms of expression within English-language culture	3
interpret literary texts and other cultural expressions from a cultural-historical and social perspective	2
elaborate on and discuss a selection of factual prose texts from English-language culture and social life	3
elaborate on and discuss the cultural position of the United States and Great Britain in the world today, and the background for the same	0
elaborate on and discuss current issues in international culture and the news media	1

In general, this table shows that only one, two or three tasks include these competence aims. It is perhaps not to expect that all would be measured in a majority of the tasks, but still, the numbers seem markedly low. Table 19 tells us that none of the tasks touch upon the topic of the cultural position of the USA and Great Britain, and only one task brings in a current issue in international culture. The topic of literary periods is also only tested once. It can be added, as well, that except for the three-paragraph task in N09 about the presidential election, all the rest of the tasks represented in table 19 are optional. The three tasks related to elaboration on factual prose are all quite short and lightweight, making it doubtful if that competence aim is really measured at all. This issue will be commented and discussed in more detail in the next chapter, but this main area seems to represent an important dimension of the construct which is under-represented in the tests.

6.6 Scoring reliability: Do they measure reliably?

The final element of this analysis regards the reliability of the scoring procedures. On the basis of the examinations and related documents, the following questions have been raised, and attempted answered:

- Are the candidates allowed too much freedom in their responses?
- Are the tasks unequivocal, and the instructions clear and explicit?
- Is there agreement on what an acceptable response is?
- Are there multiple, independent raters?

All of these questions address issues that may jeopardize the reliability of the scores.

Restricting freedom in the responses enhances reliability. It may be impossible to state precisely when too much freedom is given, but the more freedom the candidates are allowed, the harder it is for the rater to score the tests reliably. Equivocal task instructions may potentially cause different raters to interpret them differently, and hence score them differently. Similarly, if the acceptable response is not clearly defined, there is also room for different interpretations, which may lead to unreliable results. In any case, when the raters are human beings, there is a chance of human error. Using multiple, independent raters is a compensation for such error, and greatly enhances the reliability measure.

6.6.1 The Finnish examinations

One notable feature of the Finnish tests is the grammar tasks that offer a quite limited freedom of response. Freedom means to have many ways to respond to the task, so when the response has to be selected among four fixed options, the freedom is minimal. This makes it easier to ensure that the scores of those tasks are consistent, however. There is no subjective human judgment involved when scoring multiple-choice tasks; it is simply a matter of following a given protocol. Still, more than twice the amount of points of the total score is assigned to the written production, where there is more freedom of response. The freedom is restricted to some extent in the writing tasks as well since the text is to be relatively short, but the topics are often so general that in terms of content the test taker is free to adapt the response according to own preferences.

The Finnish tests follow a strict pattern, using the same type of tasks and the same instructions every time. These instructions are always short, concise and unequivocal, and describe the task as well as how the answer should be marked. There is no unnecessary information given, and the information is fairly explicit and easy to understand, leaving little room for confusion and misunderstandings. Such a well-defined framework is part of enhancing reliability.

Further, it has been shown in a previous section that the assessment documents that regulate the scoring criteria, is a good framework. The documents state that communicating a clear message is the most essential, and if it does not communicate, it is not acceptable. In addition to that, the candidate must follow the task instruction and stay on topic, as well as keeping within the given frame of length. This provides a good basis for agreement on the accepted response.

Finally, concerning the rater situation, the Finnish examination papers are first scored by the teacher at the local school, and then in an examination board by an external rater. If the difference between the teacher and the rater is too big, they resort to a second rater from the board. A difference of seven points or more in the written composition necessitates an extra rater (K. Pohjala, personal communication, November 16, 2010). Thus, there are multiple independent raters, but only one of them is external.

6.6.2 The Norwegian examinations

More freedom is allowed in the Norwegian tests, since it does not make use of any selected responses. Task 1 and 2 most often define a suitable length for the response, but there is no such specification in task 3 in any test. On the other hand, the task instructions are quite specific, restricting the response in terms of genre and topic.

All in all, the Norwegian task instructions seem fairly clear and explicit. They are not always simple and straightforward, though, which means that they risk being imprecise. An example of an expression I have found to cause some unnecessary confusion is: “Point out at least three different linguistic devices used by the author to argue his point” (Appendix 6, p.6). The confusion relates to the meaning of “at least”. What is really required from the candidate in this case? Are three examples sufficient for a full score? If so, it could have simply read: “Find three linguistic devices”. If four is better than three, it could have been specified.

However, there are also assessment guidelines meant to avoid such issues. A general assessment document as well as specific guidelines define what is to be expected for every individual task. In other words, there is agreement on what an accepted response is. The content of these documents have been mentioned already, and the main message is that the holistic picture of the entire examination paper decides whether it is acceptable or not. Moreover, it makes very explicit comments on how each task can and should be solved, which is quite necessary because of the variation of tasks from test to test. The problem, if any, is not with the specifications given, but rather with what is not mentioned. The guidelines lack information on how to judge an assignment which is notably shorter than the average. Neither are there any instructions on how to assess the whole picture if one subtask is left undone, as is the case in the Finnish documents.

As a compensation for the complexity of the instructions and rater guidelines, the Norwegian system demands two external raters for every test response. According to Norwegian law, a centrally given national examination is to be read and assessed by two external raters, and if they do not come to agreement, a third person will be called for to make the final decision (“FOR 2006-06-23 nr 724: Forskrift til opplæringslova,” 2006).

6.6.3 Comparison of scoring reliability

According to the results of the three first questions, the Finnish tests seem to be in less danger of jeopardizing reliability than the Norwegian ones. They allow less freedom of response,

provide more clarity of instructions, and have in some way a more defined agreement on acceptable response. The task instructions and assessment guidelines are more elaborate for the Norwegian tests, which increases the risk of inadequate or inaccurate information. The last question, however, points in favor of the Norwegian system's reliability since they operate with two raters that both are external.

This chapter has presented the two examination formats in terms of what they test, and in terms of possible threats to construct validity. What they test or measure has been shown by using the cognitive writing model to analyze each task. Threats to construct validity have been shown by investigating the tests' relationship with the construct, and some scoring reliability issues. The following chapter will open with a summary of the findings that have been presented here.

7 Discussion

7.1 Chapter overview

The aim of this chapter is to summarize and discuss the findings of the present study. The chapter will start with the summary, which also will relate the findings to the research statement. After the summary, I will present some reflections on the validity of these findings, before I discuss how they relate to theory. I set out to compare how the English examinations in Norway and Finland approached writing assessment, and how the two different test formats stand, in terms of construct validity. I will therefore first discuss the different approaches to writing assessment that the Norwegian and Finnish examination systems represent. Next I will discuss the construct validity of the examinations in light of what I have found. The summary of the findings below will also follow that structure, summing up the differences first, and then the findings related to construct validity.

7.2 Summary of the findings

7.2.1 The different approaches to writing assessment

Despite the similar purposes and the similar groups of examinees, I have found four major ways in which the approach to writing assessment differs notably between the Finnish and the Norwegian English examinations. That is, there are four areas in which the underlying writing constructs of the two formats are distinctly different.

First, they differ with regard to the role of reading in writing. The Norwegian test format considers reading an integrated and central part of writing, and writing tasks without reading involved is essentially non-existent. Reading comprehension is therefore measured to a great extent, including reading of relatively advanced literature. Compared to the Norwegian format, hardly any reading is required in the writing part of the Finnish examinations. Second, they differ in terms of composition topics. The Finnish tests ask for texts about personal opinions and everyday topics. The Norwegian topics deal mostly with textual analysis, and at

times with cultural or social issues. Furthermore, I found that only a surprisingly small number of tasks really require cultural knowledge. In fact, even in these tasks, the need for such topical knowledge is minimal if preparations are done well, since notes are allowed during the test. Third, they differ with regard to the specifications of genres and task settings. The Finnish format includes a few tasks where the genre and setting is specified, but adherence to genres does not have an effect on the outcome of the test score. The Norwegian tests, on the other hand, consistently ask for genre-specific texts in the longer composition tasks, and also reward the use of genre features. However, communication situations are not provided, contrary to some of the Finnish tasks.

The fourth point is the difference of cognitive level of writing tasks, and also sums up the essence of the other three differences. The Norwegian system considers writing to be the ability to draw on other written sources, conform to a given genre and to be able to deal with cultural and literary topics. The fact that the examinees of the Norwegian tests will have to relate to source material, genre specifications and sometimes specialized topical knowledge make the writing task cognitively more demanding. What is more, a command of argumentative discourse and analysis is also required in a majority of the Norwegian tasks. In comparison, the Finnish proficiency test only requires the cognitive capacity to be able to clearly and coherently explain a general or personal topic, without genre restrictions, and completely unrelated to other sources. These different approaches to writing assessment can be defended if they correspond to the respective writing constructs, but as the next section concludes, that is not necessarily the case.

7.2.2 The threats to construct validity

In both examination formats, I found signs of mismatch between the construct and the tests, threatening their construct validity. I looked at construct irrelevance, construct under-representation and reliability issues, and the threats were found mostly with regard to construct under-representation. The findings of each category are summarized below.

Regarding construct irrelevance, the overall impression was satisfactory for both examination

formats. The advanced reading, genre knowledge and high-level cognitive skills that are tested in the Norwegian examinations, seemed to be sufficiently reflected by the Norwegian construct - as defined by the syllabus. Similarly, the levels of reading, topic knowledge and cognitive skills required for the Finnish tests corresponded well with the Finnish construct - again, as defined by the syllabus. Some minor concerns are worth mentioning, however, and they will be summed up below.

Concerning the Norwegian examinations, I found two specific topical requirements in the Norwegian tasks that are not explicitly mentioned in the construct, knowledge of the Romantic literary period and knowledge of picture analysis. The construct only includes these elements in general descriptions, and the nature of these general descriptions make it difficult to say with certainty whether they are included. The issue for the Finnish tests is the mismatch between test construct and scoring criteria. Situation appropriate language is clearly a test requirement according to the criteria, but the construct mentions nothing of the sort. One may add that, in this area, there is also a mismatch between the scoring criteria and the task instructions, since only a few of the tasks give the examinee an opportunity to show situation appropriate language.

The findings involving construct under-representation were more revealing, indicating potentially critical threats to construct validity, perhaps especially, but not exclusively, for the Norwegian format. The Norwegian construct comprises three main areas, each of which include several competence aims, and one of these areas is nearly completely neglected in the sample of tests studied here. The Norwegian tests are to measure the content of the area called “culture, society and literature”, but, in my opinion, this is not done to a sufficient degree in the examinations I have examined.

Significant results were also found related to construct under-representation in the Finnish tests. The construct of the Finnish format is fairly simple, covering the basic skills related to writing proficiency. Still, there were two components of the Finnish construct that were not followed sufficiently up by the tests. Knowledge of basic social and formal genres is one of them. However, only five out of twelve composition tasks are genre-specific, and more importantly, genre knowledge is not part of the scoring criteria. The other construct

component, which has not been tested in any of the tasks, is the ability to summarize information from different sources in a text.

Last, I will summarize the findings concerning reliability issues in relation to the overall construct validity. Even if the tests are perfect representations of their constructs, one cannot justify the use of the test scores if rating is not consistent. In writing assessment, there is often a chance that subjective judgment could cause inconsistent scoring. This chance decreases if the tests are carefully constructed to avoid misunderstanding and unnecessary variation, and if the same care is taken to clarify the scoring criteria and accepted responses. Essentially, I found both tests to have taken such care to a reasonable degree. In both formats, there are sets of instructions for scoring and accepted responses, as well as mostly unambiguous task instructions as well. The Norwegian tests, since they include more detailed information than do the Finnish ones, seem to be in greater danger of opening up for unwanted influence on test scores. However, both nations have also taken measures to prevent too much subjective influence, using a system of two, and if necessary three, raters for each test. I cannot assert for certain that the reliability is high, but neither have I found any critical issues. However, this would belong to a separate study of inter-rater reliability that could probably be done using rating information available in the databases of the Norwegian Directorate of Education and Training.

7.3 Validity of the findings

With regard to the validity of my study, it is to emphasize that the results are only valid for the tests I analyzed, due to my limited sample with only three tests from each country. This does not mean that the findings are uninteresting in a larger perspective, but it means that it is not possible to transfer these results to describe the Finnish and Norwegian English examination system as such. However, if time had permitted the inclusion of all of the nine available Norwegian tests for the literature course, for instance, this would have allowed firmer conclusions about the Norwegian system.

Moreover, despite my attempts to be as systematic and clear as possible, I have had to make some decisions along the way of what to include and what not to include, to keep the thesis precise and readable. My choice of using the cognitive model to structure the investigation of the examinations tasks, I believe, was part of improving the systematic approach and validity. However, the decisions made when categorizing genre specifications and types of writing were less systematic and might therefore be more discussable. This means that even for the tests that have been analyzed, one has to be somewhat cautious when drawing conclusions from the results. Yet, I will argue that the main findings are largely unaffected by the details of these decisions, and that the results to a reasonable degree give a useful picture of the tests in the sample.

Finally, I will also comment on the findings related to scoring reliability. The reliability questions I used are directly from a list by Johnson (2001), and should therefore be well validated. I also believe that my findings give some indication about the reliability of the tests in question. However, I feel that I lacked a systematic procedure to ensure that every side of the issue was investigated. I also lacked a set of criteria by which to judge whether the task instructions and assessment guidelines were clear and unambiguous or not. In the section on further research in chapter 8, I therefore propose another, and arguably better, way to examine the reliability issues.

7.4 Discussing the different approaches to writing assessment

7.4.1 Testing academic writing

In chapter 3, we looked at the challenges of defining writing ability as a skill. In fact, Bachman and Palmer (2010) claimed that each writing activity is different, and that each one should be looked at as a separate construct. Therefore the chapter listed different types of writing that are useful for different types of writers. Indeed, one of the differences between the Norwegian and Finnish tests is the different writing activities they elicit. As mentioned above, writing in the Norwegian tests comprises reading, referring to sources, discussing,

including relevant content and conforming to a given genre. The cognitive demand is clearly higher than in the Finnish tests, where writing consists of producing 200 words of consistent, coherent text about personal opinions. In fact, it resembles the difference between what Bereiter and Scardamalia (1987) termed knowledge telling and knowledge transformation. As described in chapter 3, knowledge telling is mainly writing from the top of your head. It is a natural process “because it makes use of readily available knowledge - thus it is favorable to report of personal experience” (Bereiter and Scardamalia, 1987, p.9). In other words, it applies well with the writing types in the Finnish test format. Knowledge transformation is a laborious problem-solving strategy for writing, where working out problems of content and rhetorics may result in gaining new knowledge in the process. The argumentative texts and source referencing involved in the Norwegian tasks seem allow for testing of this type of writing. Are these two approaches to writing assessment equally justified? Why is writing ability defined so differently in these constructs?

The basis for developing a syllabus or a construct is often the needs of the language learners (Simensen, 2007). The question is therefore what type of writing the learners of English need to be proficient in at this level. Is there need for more than story telling in a foreign language? Weigle (2002) pointed out that there is, “for students nearing the end of compulsory education” (p.11). According to Weigle (2002), there is a need for knowledge transformation skills at that level, since it is related to academic writing skills. It is arguably not the same need for academic writing in a foreign language as in the mother tongue. It is therefore interesting to notice that, contrary to the Norwegian curriculum, the Finnish curriculum defines English only as one of many foreign languages. However, I would argue that English is in a special position, and that academic writing in English is relevant for higher education in the Nordic nations.

Academic writing can perhaps not be expected in upper secondary school, but as Sparboe (2008) points out in his study, upper secondary school should still prepare the students for higher education. He found that this was not the case for the R94 syllabus (Sparboe, 2008), but based on my findings, I would argue that the present examinations in Norway do include aspects of academic writing. The Finnish construct, however, limited to a knowledge-telling

approach to writing, is perhaps not that well adapted to the future needs of the learners at this level.

It is possible that the Finnish examination system has prioritized tests that are uncomplicated and quick to assess reliably for the sake of practical reasons or cost-effectiveness. Still, while it is important that a test measures what the construct describes, it is not indifferent what type of construct is being measured either. In fact, in my pilot study, one of my checklist questions from Bachman (1996) illuminates this point: “Is the language ability construct for the test relevant to the purpose of the test?” (p.141). For the Finnish matriculation examinations, this issue is arguably more critical than whether the test conforms to the construct. In the following section, I will continue to discuss the basis for the writing constructs used in the tests.

7.4.2 Testing communicative competence

Testers need to be applied linguists, aware of the latest and most accepted theories of language description, of language acquisition and language use. They also need to know how these can be operationalised: how they can be turned into ways of eliciting a person’s language and language use (Alderson & Banerjee, 2002, p. 80).

Test constructs should not only be based on the needs of the learners and society, but should also reflect the most recent research (Alderson & Banerjee, 2002). Recent research in language theory describes language as communicative competence, and the national curricula of both nations are based on the CEFR, and its concept of communicative competence. However, in view of the findings of this study, I find it necessary to discuss and shed some light on the testing of communicative competence.

The essence of communicative competence is that language is more than decontextualized rules of grammar and syntax. Language is a means of communication used to serve a purpose in a social context (Council of Europe, 2001). How can such a view of language be

adequately tested, and what does it entail for the examinations in question? Certainly, both test formats include more than grammar in the assessment, but do they test language as communicative competence when there is only minor focus on social contexts?

It seems to me that the two different approaches to writing allow for different degrees of testing the socio-linguistic aspect of language. The dilemma seems to be whether to include tasks involving a social context, as some of the Finnish tasks do, or to include tasks of more academic genres, as in the Norwegian format. The challenge for the Norwegian format is that the academic genres are often incompatible with real-world social contexts. What would be the intended audience and purpose of poetry analysis, except to show the teacher your academic skills? There is no authentic communicative setting involved in such a task, it only relates to life in academia. However, these types of tasks allow for testing of knowledge transformation as explained above. Not all the Finnish task types include a reference to real life contexts, but those that do easily become knowledge telling exercises.

If it is a choice between testing knowledge transformation and the ability to adapt language to the social context, I believe it has to be made based on the purpose of the test. Based on the previous section, it seems reasonable that testing knowledge transformation is of greater importance in these types of tests, where the purpose, at least in part, is preparation for higher education. Bachman and Palmer (2010) also point out that every language assessment does not have to encompass the whole range of communicative competence, but they need to be informed by a broad view of language. Another difference between the two writing approaches will be dealt with next.

7.4.3 Testing topical knowledge

Weigle (2002) refers to an ongoing discussion on how to deal with topical knowledge in writing assessment. Bachman (1996) speaks of three possible ways of dealing with it in a test: to avoid topical knowledge, to include it in the language ability construct, or to include it in a separate construct. In my analysis, we have seen two examples of these variations. The Finnish examinations avoid topical knowledge, being limited to general and personal topics.

In the Norwegian construct, topical knowledge is most certainly included, and the content of the writing tasks is also sometimes provided in appended reading material.

“Perhaps the most important consideration here is accessibility to all test takers, since everyone needs to have an equal chance of success” (Weigle, 2002, p.91). Using reading material as basis for the written composition, as in the Norwegian approach, is useful since it provides all the test takers with the same content information, giving everyone a fair change. However, it also poses a challenge for poor readers. Poor readers may suffer if they misinterpret the text, and may be put at disadvantage regardless of their writing abilities (Weigle, 2002). On the other hand, the challenge involved in using general topics, as in the Finnish approach, is creating engaging topics for everyone. Writing about personal experience seems like a good solution, but one has to keep the balance between making topics that engage the test takers, while avoiding too emotional responses (Weigle, 2002). A personal topic may, in one extreme, elicit such an emotional response that the language and structure suffer, while a non-personal topic, on the other hand, may fail to engage the interest to write, and elicit mechanic and uninteresting compositions (Weigle, 2002).

When including topical knowledge in the construct, these issues can be avoided, but it means that it is no longer only language skills which are being assessed. The question is then whether cultural and social content belong in language assessment. Simensen (2007) claims that “it is virtually impossible to teach a language without teaching cultural content” (p.79). Is the same true for testing languages? Can or should culture and language be separated? Even if cultural topics are to be included, there is still a question of which cultures to consider, especially concerning the English language. The Norwegian test construct (see section 4.2.3), for instance, refers mainly to American and British culture, but one competence aim also mentions “current issues in international culture” (Udir, 2006, p.7). English, being an internationally used language, makes it difficult to narrow the scope of cultural interest.

Nevertheless, if topical knowledge is included in the construct, it should also be tested, and vice versa. That will be one of the subjects in the next section when discussing construct validity, and particularly in section 7.5.3, titled “construct mismatch”.

7.5 Discussing construct validity

7.5.1 Brief review of construct validity and validation

The research statement for this study is: *“How does the construct validity of writing assessment compare between Finnish and Norwegian English examinations at the highest level of upper secondary school?”* Construct validity was defined in chapter 2, not as a test quality, but as “the extent to which test interpretations and uses can be justified” (Chapelle, 1999, p.258). The justification process is what was called validation, and means gathering evidence to convince the people affected by the test that this is the case (Bachman & Palmer, 2010). Since construct validity is not the quality of a test, it is impossible to say that any of these tests in themselves lack construct validity. The question is whether it is possible to justify the use of the test scores, and find evidence that support it.

In order to justify the use of the scores, one has to find evidence that the scores represent the test takers' knowledge of English as defined in the syllabi. I have examined to see if there is evidence showing something else. Can we claim that the scores correctly reflect the test takers' performance on the test? And can we claim that the performances give a picture of the abilities as the construct defines it? Based on the findings, the two next subsections will elaborate on and discuss these two questions.

7.5.2 Reliability

If the test scores are not reliable, they cannot be said to represent the test takers' performances on the test. Unreliable scores mean that the scores are also influenced by other factors, such as the rater's understanding of the scoring criteria, or his or her subjective judgment.

In language tests using open-ended writing tasks, there will never be a clear-cut right or wrong answer, and they are therefore bound to include some form of subjective judgment. In other words, we cannot know for sure that the scores of these tests would have been the same if scored by another rater, for instance. Some uncertainty about the reliability is unavoidable if we open up for the freedom of responding to open questions, as is the case for the

examinations in this study. However, the findings show that the raters are able to follow carefully written assessment guidelines, and that the task instructions are written to avoid ambiguity. Even if there are imperfections, and there is subjective judgment involved, there are always two people rating the tests, which compensates for human error.

The tasks could have been created to avoid subjective judgments, but this would have been at the cost of something else. By only using multiple-choice tasks, the scoring could come close to complete reliability, since human judgment is minimized. In this respect, the Finnish tests are arguably more reliable, one part of the score being based on multiple-choice tasks. However, increased reliability often comes at the cost of the overall construct validity. To put it simply, how would writing ability be tested, if the test takers were not to write anything? If a writing construct were to be tested only by multiple-choice tasks, we would not be able to measure these components from the Finnish and Norwegian writing constructs:

- produce texts in a variety of genres with clear content, appropriate style, good structure, and usage that is precise and accurate (Udir, 2006, p.6).
- Can use broad vocabulary and demanding sentence structures together with linguistic means to produce a clear, cohesive text (Finnish National Board of Examination, 2004, p. 247).

Nevertheless, based on the analysis, I have found little reason to call into question the reliability of the scoring procedures of the tests in question. In order to convince stakeholders that the scores represent the skills of the testees, more research would be necessary, but I have not been able to show this as a weakness to the overall construct validity. Assuming that the scores are reliable, there is still a question of whether they reflect the test takers' skills as defined in the construct?

7.5.3 Construct mismatch

Test takers' performance is elicited by the test tasks, but do these tasks reflect what the construct describes? Based on the findings of this study, are we able to convince the test takers and society that the tests faithfully reflect what is defined in the syllabus? Can we

convince them that it does not test anything outside the construct, and that what they test fairly represents all the major aspects of it?

Construct irrelevance

I believe it is possible to convince the people involved in this assessment that none of these examinations measure anything outside the construct. Despite some signs indicating that there is not a full parallel in this matter, it is possible to defend the tests against these threats to a reasonable degree.

In the Finnish tests, I found that situation appropriate language was tested, but not included in the construct. If one could argue that it is included in the construct, or that it is not tested, one would discharge this construct validity threat. First of all, not all the tasks measure situation appropriateness since they do not specify a situation. Second of all, one could argue that it is included in the construct when it states that the pupils “Can write (...) routine factual messages and more formal social messages” (Finnish National Board of Education, 2004, p. 247). One may claim that the construct covers it, since knowing how to write a formal social message might include knowing how to adapt the language to the social context as well. On the other hand, I would perhaps argue that it is possible to master a genre, and use the correct genre features, but still fail to adapt the language to the situation. Consider the following task: “You will be participating in European Youth Week and will give a speech on a topic you find important to all young Europeans. Write this speech” (Appendix 2, p.18). The writer may know the speech genre, and be able to construct a formal speech, but still be insensible to the situation, and fail to adapt the language to young people. I have therefore defined this issue as construct irrelevance, but it is not the most obvious and critical case, and one may be convinced otherwise as well.

I also believe it is possible, and not so challenging, to convince the stakeholders that the Norwegian tests do not measure anything outside the construct. I have mentioned that one task requires knowledge of the Romantic literary period, and that the construct does not specifically include this period. Again, if it is possible to argue that the construct actually includes it, or that the test does not require it, one would eliminate the threat. First, this task is

only one of four options to choose from, which means that it is not strictly required. Second, it is reasonable to accept that a representative selection of literary periods, as the construct puts it, would naturally include the Romantic period. In other words, it does seem like the issue of construct irrelevance weakens the construct validity of these tests.

Construct under-representation

While the threats caused by construct irrelevance were minor and easy to remedy, I find it rather doubtful whether all the construct under-representation issues can be explained away. Being aware of the findings of this study, how can anyone argue that both examination formats are fully representative of their constructs?

The results from the Finnish tests show that only five out of twelve composition tasks include genre specifications, despite the fact that the construct includes the genre aspect. The word “genre” is not used in the construct, but the ability to write “routine factual messages and more formal social messages” is mentioned, and exemplified by genres such as “reviews, business letters, instructions, applications [and] summaries” (Finnish National Board of Education, 2004, p. 247). Therefore, there seems to be no doubt that a degree of genre knowledge is part of the construct.

The question is therefore whether this aspect is sufficiently tested in the examination tasks. Genre specifications are not absent from the Finnish tests, but it is possible to avoid genre-specific tasks in every test. It is also possible to choose it, and perhaps one may argue that the tests do not fail to include the genre aspect just because it is optional to write in a set genre. However, the fact that the genre aspect is not even included in the scoring criteria, settles the case. Since the scoring criteria do not take genre into account, it is certainly not tested, and there is an apparent mismatch with the construct. If someone fails to write the specified genre in a task, but otherwise writes in accordance with the criteria, he or she should be able to receive the top score. Consequently, I would argue that this is an important dimension of the Finnish construct which is not adequately represented in the tests. Another case of under-representation was the complete lack of tasks requiring summarizing of different sources. This is clearly stated in the construct, but clearly not asked for in any task, so the mismatch

seems obvious. The question is perhaps whether this is an important dimension which is left out, or whether it is a rather minor part of the proficiency scale. To me, it seems to represent an aspect of some importance, but perhaps not as important as the genre aspect. On the other hand, referring to other sources is one of the characteristics of academic writing that we also found missing in the Finnish test format above.

The situation does not seem to be any better for the Norwegian examinations with regard to construct under-representation. I have found two possible threats, and only one of which can be remedied. In the main area of communication in the Norwegian construct, there was one component which was markedly less represented than the others, which was to “summarize, comment on and discuss *differing viewpoints* in fictional texts” (Udir, 2006, p.6). First, one has to notice that this is a fairly specific competence aim. In other words, one cannot say that it represents an important dimension of the Norwegian construct. Moreover, although it is only tested in three tasks, all of the three tests include such a task, indicating that it is fairly well represented after all. It should therefore be easy to convince the stakeholders that the major aspects of the communication aims are adequately represented in the tests. Convincing them that the area of culture, society and literature is just as well represented may, however, be more demanding.

There is no doubt that the subject main area *Culture, Society and Literature* represents an important dimension of the Norwegian construct. There are seven competence aims in this area, which counts for almost half of all the competence aims for this course. Three of those seven competence aims are tested in only three tasks, while the remaining four are tested in less than three. It is understandable that these numbers are not as high as for the communication area, and it may also be acceptable that one of the elements is not measured at all, since the sample of tests is relatively small. However, when all the numbers are so low, the situation seems critical.

Indeed, not only are the numbers low, but they could have been lower. I have been liberal when counting tasks that measure the following competence aim: “elaborate on and discuss a selection of factual prose texts from English-language culture and social life” (Udir, 2006, p.6). Table 19 showed that three tasks measure this ability, but here is an example of such a

task: “Write a paragraph in which you comment on content, style and language features that indicate that this is an academic text” (Appendix 5, p.6). The academic text referred to is a short passage of thirteen lines. In my opinion, writing one paragraph commenting on content and language of a text is hardly a way to show elaboration and discussion of factual prose. If this type of tasks would not count as testing ability to discuss factual prose, there would not be any of them in the sample tests.

Finally, as has been mentioned in the summary above, one may question whether any of these competence aims with regard to topical knowledge are at all relevant, considering that students are permitted to bring books and notes to the test. If the candidates would only prepare well, they could easily bring a list of genre features for all the literary-historical periods since the Renaissance. In such a case, the task does no longer seem to measure competence in the area of culture, society and literature, but the ability to find and use sources. In conclusion, therefore, I see no other way than to maintain that this area is, at best, clearly under-represented in the sample tests, and hence poses a serious threat to the construct validity of the Norwegian examinations.

7.6 Test consequences

Do these findings really matter? Is it really of any importance how the assessment is conducted? If these tests have been administered for years already, and seem to work well, is it not enough? These are legitimate and significant questions to ask, but we need to bear in mind that the use of assessment has consequences (Bachman & Palmer, 2010), as stated in chapter 2. Validity matters if the consequences matter. In this case, we are talking about very high-stakes tests with consequences for a tremendous amount of young people every year. The score of these tests will have direct influence on these people's access to higher education, and thus even affect their professional choices. These decisions should not be made on the wrong grounds. When the stakes are as high as this, and the tests are administered at the national level, mistakes cannot be accepted.

Test consequences have not been the focus of this study, but in order to emphasize the importance of developing valid English examinations, I will give a final comment about the

phenomenon of washback. Messick (1996) defines it as “the extent to which the introduction and use of a test influences language teachers and learners to do things they would not otherwise do that promote or inhibit language learning” (p.241). It deals with consequences of tests on teaching and learning.

In Finland, one study by Pietilä, Taanila-Hall and Vainio (2009), which was mentioned in the introduction, concluded that the English matriculation examination has influenced use of speaking exercises in textbooks. The other study on washback by Vainio (2009) did not find significant evidence of textbook washback, but still concluded as follows:

Although there is not much washback at the textbook level, in the classroom the washback effect is probably visible. Especially in the last courses the matriculation examination is likely to guide the actions of both teachers and students (Vainio, 2009, p.110).

The Norwegian washback related study by Ellingsund (2009), analyzing Vg1 English examinations, found “no causal relationship between innovations in testing and innovations in teaching”, and that “the influence of a test has different outcomes for individual teachers” (Ellingsund, 2009, p.126). Nevertheless, seven out of twelve teachers confirmed that the test influenced their teaching to some degree, and four of them judged it to be highly influential in their teaching (Ellingsund, 2009).

In other words, examinations may sometimes influence textbooks and teaching. A high-stakes examination is likely to have quite much influence, so making sure that they meet the needs of the learners, are based on recent research, are scored reliably and correspond to the construct is of utmost importance.

7.7 Discussion summary

This chapter has summed up the findings of the study in two sections. One part summarized the differences of the approach to writing assessment in the two examination formats, and the other summarized the findings related to the construct validity of the tests. After the summary

of findings, I also discussed their validity, emphasizing the limitations related to the small sample size and subjective decisions.

The different approaches to writing and construct validity of the tests were also the two parts that were discussed. With regard to the differences, I discussed the challenge of defining a writing construct for examinations in upper secondary school. Considerations that were discussed included whether or not, and how, to incorporate aspects of academic writing, specialized topical knowledge and socio-linguistic competence. I also discussed the basis for such decisions, such as the needs of the learner and society, recent language research and the purpose of the assessment. In this discussion, I argue for the importance of a construct that is useful for its purpose. In that respect, I questioned the lack of focus on academic writing skills in the Finnish examinations.

The other discussion part attempted to weigh the evidence for or against construct validity. The focus was on scoring reliability and the relationship between test and construct. The reliability issue was not found to be a threat to construct validity in any of the tests, but more research in this field is required to say anything conclusive. Evidence of construct irrelevance were discussed, but refuted as it seemed possible to explain it away. The evidence of construct-underrepresentation, however, was in both test formats found to be hard to counter, meaning that it represents a real threat to their construct validity.

Finally, I briefly discussed whether any of these issues really matter, arguing that they do, not only because it affects students' admission to higher education, but also because high-stakes tests are likely to affect textbooks and teaching instruction. In the next and closing chapter, some implications of these findings will be mentioned, as well as suggestions for further research of Nordic English examinations.

8 Conclusion

High-stakes nationally administered final examinations of two Nordic countries have been investigated with regard to their construct validity and their differences. I wanted to know how the characteristics of the test formats were, in comparison to one another, and how they were in relation to their respective English syllabi. The comparison was of interest because the tests serve quite similar purposes for a quite similar group of upper secondary school students.

First, the analysis first revealed two completely different views on writing reflected in these tests, where the general difference is that the Norwegian examinations test considerably more academic writing features than the Finnish ones.

Second, for both test formats it was found that the writing constructs included more elements than what the tests measure. A great amount of the topical knowledge emphasized in the Norwegian syllabus was not adequately tested by the Norwegian tests. And despite the Finnish English syllabus making mention of genre knowledge and summarizing, none of it was really tested in the sample of Finnish examinations. In other words, both test formats fail, to some extent, to adequately measure their constructs, which poses a real threat to their construct validity. It means that those responsible for developing and using these examinations in Norway and Finland will have a challenge justifying the use of test scores, since they do not fully represent the constructs outlined in the curricula. In this concluding chapter, I will bring the thesis to a close by proposing further research in this area, and by mentioning some possible implications of what I have found.

8.1 Further research

Since the findings are limited to the six examinations in the sample, determining whether or not this study shows a tendency for the examination systems in general would require more research. Based on the potentially critical findings of this study, I believe it would be worth

spending more time and effort investigating the national examination systems of Norway, as well as other Nordic countries, with special focus on construct validity.

With regard to the Norwegian situation, I believe it would be useful to further assess the quality of the relatively recent Vg3 level English courses and examinations. An option could be to expand the analysis of this study, and include all the Norwegian examinations related to the English course of culture and literature. In this way, one would be able to provide more reliable information about the total situation of these examinations, and could concentrate more on the specific issues relevant to Norwegian English assessment.

Nothing critical was found in this study regarding the reliability of the test formats. My findings in this area, however, were partly inconclusive due to lack of actual scoring data. Looking back at the thesis now, I see that a more thorough analysis of test reliability could have strengthened the quality of the study. To investigate this area in further detail, I would propose a study of rater behavior, using a sample of around ten test responses from each nation, and their respective scores. A set of qualified raters, unaware of the original score, would rate them two by two according to the respective assessment guidelines. A comparison of the original scores and the new scores from the selected raters would supply useful data on reliability issues, including the quality of the test instructions and rater guidelines, as well as on the effect of the two-rater system.

Moreover, there is available data at the Norwegian Directorate of Training and Education to examine inter-rater reliability in the last years' English examinations, as well as for other subjects, in a large-scale, quantitative study. Since the two raters' proposed grades and the final examination grade are available in the Directorates databases, it would be possible to analyze the difference between the two. The average difference in a large enough sample could be a significant indicator of level of inter-rater reliability. A high correlation (Pearson's r) would mean high inter-rater reliability, and vice versa.

Last, it would have been useful to combine the content analysis with real think-aloud protocols. That is, to have students, who actually sit the tests, explain out loud how they think when they solve the tasks. Such a study could confirm or disconfirm the assumptions made in this study about what the tasks require. It would provide additional information about what the tasks really measure by learning how the test takers approach the tasks, and what skills they actually employ. My hypothesis is that many elements would appear which are not taken into account in this study. A more accurate picture of what skills the tests involve would also improve the investigation of construct irrelevance and construct under-representation.

8.2 Implications

To the extent that the results of this study say anything valid about the tendencies in the national examination systems as a whole, I would argue that there are several implications of these findings for Norwegian as well as Finnish school authorities.

For test and syllabus developers of the Norwegian examination format, I believe this study gives reasons to evaluate the role of cultural knowledge in the tests. If the subject main area of culture, society and literature is to be properly tested in the examinations, the essay tasks should actually require the use of such knowledge. It is imperative to ensure that the tests are adjusted in the future to also measure the topical aspects adequately. Another option would be to remove the topical aspect from the test construct. Reading material could still be used as a basis for topic content, and thus continue to test the ability to discuss cultural and social issues.

What is more, I believe that Norway could learn from the Finnish format in a couple of areas. In the Norwegian examinations, the writing tasks do not allow for writing to be tested independently of reading, as the Finnish tasks do. This could be solved if only one of the three writing tasks were made without reference to a text. Testing of reading comprehension could instead be strengthened by implementing a separate reading comprehension part, similar to the first part of the Finnish test.

For the situation in Finland, it seems most crucial to consider the level of academic writing required for the matriculation examination in English. If it is meant to prepare the students for higher education, is necessary need to involve more expository and argumentative texts. This might require another test format, where the student can write a longer text, like in Norway, and where genre knowledge and perhaps source referencing is incorporated. Such a change would probably entail a different and more costly scoring system, but I suggest that the option should be considered. In any case, the findings at least suggest that the Finnish examination board would also take some measures to adjust their tests to fit the construct, or alternatively alter the construct definition. One specific and simple proposal, which would strengthen the construct validity immediately, would be to include genre knowledge as one of the scoring criteria.

Lastly, the findings also suggest a need for closer cooperation between the Nordic countries, and perhaps a plan of how to even out the differences of English writing in upper secondary school. This is especially relevant now since it was recently decided that university applicants from the Nordic countries will be given the same chances of admission to higher education regardless of whether you apply in your own country or another one of the Nordic nations (NTB, 2012).

It is likely that high stakes educational assessment will continue to play an important role in our society for a long time. And there will always be uncertainty involved when making the inference from what the students know to a simple number on a piece of paper. However, although we cannot expect perfection, it is always possible to improve, and we must never cease to put the test to the test.

References

- Act of 17 July 1998 no. 61 relating to Primary and Secondary Education and Training (the Education Act), § 3-4 (1998). Retrieved from <http://www.ub.uio.no/ujur/ulovdata/lov-19980717-061-eng.pdf>
- Alderson, J. C., & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, 35(02). doi:10.1017/S0261444802001751
- Ary, D., Jacobs, L. C., & Sorensen, C. (2010). *Introduction to Research in Education*. Wadsworth, Canada: Wadsworth Cengage Learning.
- Bachman, L. F., Davidson, F., & Milanovic, M. (1996). The Use of Test Method Characteristics in the Content Analysis and Design of EFL Proficiency Tests. *Language Testing*, 13(2), 125–150. doi:10.1177/026553229601300201
- Bachman, L. F., & Palmer, A. (2010). *Language Assessment in Practice*. Oxford, England: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford, England: Oxford University Press.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Broadfoot, P. (2007). *An introduction to assessment*. New York, NY: Continuum.
- Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.
- Chapelle, C. A. (1999). Validity in Language Assessment. *Annual Review of Applied Linguistics*, 19, 254–272. doi:10.1017/S0267190599190135
- Cohen, A. D. (2006). The Coming of Age of Research on Test-Taking Strategies. *Language Assessment Quarterly*, 3(4), 307–331. doi:10.1080/15434300701333129
- Council of Europe. (2001). *Common European framework of reference for languages: learning, teaching, assessment*. Retrieved from http://www.coe.int/t/dg4/Linguistic/Source/Framework_EN.pdf
- Cumming, A., Grant, L., Mulcahy-Ernt, P., & Powers, D. E. (2004). A teacher-verification study of speaking and writing prototype tasks for a new TOEFL. *Language Testing*, 21(2), 107–145. doi:10.1191/0265532204lt278oa

- Ellingsund, A. E. (2009). *Washback in EFL instruction in Norwegian upper secondary school: has the new ENG1002/03 exam led to positive washback effects in EFL instruction on the English foundation course?* (Master), Norwegian University of Science and Technology, NTNU, Trondheim.
- Finnish National Board of Education. (2004). *National core curriculum for upper secondary schools 2003: national core curriculum for general upper secondary education intended for young people*. Vammala, Finland: Finnish National Board of Education.
- FOR 2006-06-23 nr 724: Forskrift til opplæringslova. (2006, June 23). Retrieved January 26, 2012, from <http://www.lovdata.no/for/sf/kd/xd-20060623-0724.html>
- Giverholt, H. (2009, March 23). Finlandssvensker. *Store norske leksikon*. Retrieved from <http://snl.no/finlandssvensker>
- Guo, L. (2011). Product and Process in Toefl iBT Independent and Integrated Writing Tasks: A Validation Study. *Applied Linguistics and English as a Second Language Dissertations*. Retrieved from http://digitalarchive.gsu.edu/alesl_diss/19
- Huhta, A., Kalaja, P., & Pitkänen-Huhta, A. (2006). Discursive construction of a high-stakes test: the many faces of a test-taker. *Language Testing*, 23(3), 326–350. doi:10.1191/0265532206lt331oa
- Johnson, K. (2001). *An introduction to foreign language learning and teaching*. Harlow, England: Longman.
- Krippendorff, K. (2012). *Content analysis*. Thousand Oaks, CA: Sage.
- Lazaraton, A., & Taylor, L. (2007). Qualitative research methods in language test development and validation. In Janna D. Fox, Mari Wesche, Doreen Bayliss, Liying Cheng, Carolyn E. Turner, & Christine Doe (Eds.), *Language Testing Reconsidered* (pp. 113–130).
- Luoma, S. E. (2001). *What does your test measure? Construct definition in language test development and validation*. Retrieved from http://www.solki.jyu.fi/vanhat/Luoma_Sari_2001_PhD_manuscript1.pdf
- Luoma, S. E. (2002). What is happening in language testing? *Sprogforum*, 8(23), 7–15.
- Marshall, C., & Rossman, G. B. (1999). *Designing Qualitative Research*. Thousand Oaks, CA: Sage.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241–256. doi:10.1177/026553229601300302
- NTB. (2012, oktober). Nordiske studenter likestilles i høyere utdanning. *Aftenposten*. Retrieved from <http://www.aftenposten.no/jobb/Nordiske-studenter-likestilles-i-hoyere-utdanning-7032690.html>

- O'Sullivan, B., Weir, C. J., & Saville, N. (2002). Using observation checklists to validate speaking-test tasks. *Language Testing*, 19(1), 33–56. doi:10.1191/0265532202lt219oa
- Organisation for Economic Co-operation and Development. (2001). *PISA Knowledge and Skills for Life First Results from PISA 2000: First Results from PISA 2000*. OECD Publishing.
- Organisation for Economic Co-operation and Development. (2004). *PISA Learning for Tomorrow's World First Results from PISA 2003: First Results from PISA 2003*. OECD Publishing.
- Organisation for Economic Co-operation and Development. (2007). *PISA PISA 2006 Science Competencies for Tomorrow's World: Volume 1: Analysis: Science Competencies for Tomorrow's World: Volume 1: Analysis*. OECD Publishing.
- Pettersen, S. (2001). *The Foundation Course in English: Some Aspects of the Written Exam*. (Master), The University of Oslo, Oslo.
- Pietilä, P., Taanila-Hall, L., & Vainio, S. (2009). Instructional writing on English: What guides the textbook writer? In M. Peikola, J. Skaffari, & S.-K. Tanskanen (Eds.), *Instructional writing in English: Studies in Honour of Risto Hiltunen* (Vol. 189, pp. 211–232). Amsterdam: John Benjamins Publishing. Retrieved from <http://metro-natshar-31-71.brain.net.pk/articles/9027254249%20Instructional%20Writing1.pdf#page=226>
- PISA. (2005, 2007). *Store norske leksikon*. Retrieved from <http://snl.no/PISA>
- Plakans, L. (2009). Discourse synthesis in integrated second language writing assessment. *Language Testing*, 26(4), 561–587. doi:10.1177/0265532209340192
- Shirazi, A. N. (2010). *English academic writing in Norwegian upper secondary school*. (Master), The University of Oslo, Oslo. Retrieved from <http://urn.nb.no/URN:NBN:no-26431>
- Simensen, A. M. (2007). *Teaching a foreign language: principles and procedures* (2nd ed.). Bergen: Fagbokforlaget.
- Sparboe, M. L. S. (2008). *English academic writing in Norwegian upper secondary school*. (Master), The University of Oslo, Oslo. Retrieved from <http://urn.nb.no/URN:NBN:no-21061>
- Studentexamensnämnden. (2007, September 21). *Språkproven föreskrifter og anvisningar*. Retrieved from http://www.ylioppilastutkinto.fi/maaraykset/yleisohje_ruotsi101.pdf
- Suomen kieltenopettajien liitto. (n.d.). Tilastotietoa kielivalinnoista. Retrieved March 9, 2012, from http://www.sukol.fi/medialle/kielivalinnat/tilastotietoa_kielivalinnoista
- Thompson, G. (2004). *Introducing Functional Grammar* (2nd ed.). London, England: Arnold.

- Utdanningsdirektoratet. (2006). *English - Programme subject in programmes for specialization in general studies* (8th ed.). Retrieved from www.udir.no/Upload/larerplaner/Fastsatte_lareplaner_for_Kunnskapsloftet/english/Languages/5/English_programme_subject.rtf
- Utdanningsdirektoratet. (2009). *Trekkordning ved eksamen i Kunnskapsløftet*. Retrieved March 23, 2012, from <http://www.udir.no/Regelverk/Rundskriv/2009/Udir-1-2009-Trekkordning-ved-eksamen-i-Kunnskapsloftet/>
- Utdanningsdirektoratet. (2011a). *Fag- og timestfordelingen i grunnopplæringen – Kunnskapsløftet*. Retrieved from http://www.udir.no/Upload/Rundskriv/2011/Udir-1-2011-Fag-og-timestfordeling_des2011.pdf?epslanguage=no
- Utdanningsdirektoratet. (2011b). *Vurderingsveiledning*. Retrieved from http://www.udir.no/Upload/Eksamen/Videregaende/V2011/Vurderingsveiledninger/5/SPR3012_Engelskspraklig_litteratur_og_kultur_BM.pdf
- Utdanningsdirektoratet. (2011c). *Sensorveiledning 31.05.2011 SPR3012 Engelskspråklig litteratur og kultur*. Retrieved from http://www.udir.no/Upload/Eksamen/Videregaende/V2011/SensorveiledningerV11/4/SPR3012_Sensorveiledning%20_BM_V11.pdf
- Utdanningsdirektoratet. (n.d.). *Udir.no - Fag: Engelsk 1.-2. årstrinn*. Retrieved March 29, 2012, from <http://www.udir.no/Lareplaner/Grep/Modul/?gmid=10&gmi=58526>
- Vainio, S. (2009). Does Multiple-Choice Testing Lead to Multiple-Choice Teaching? Textbook washback of the Finnish matriculation examination in English. In L. Kokkola (Ed.), *Tracks to Mastering English* (pp. 94–107). Turku, Finland: Anglicana Turkuensia. Retrieved from <http://www.doria.fi/handle/10024/45385>
- Weigle, S. C. (2002). *Assessing writing*. Cambridge Language Assessment Series. Cambridge, England: Cambridge University Press.
- Westermarck, R. (Ed.). (2011). *Utbildningsguiden 2010 Svenskspråkig utbildning för unga och vuxna*. Retrieved from <http://www02.oph.fi/koulutusoppaat/Utbildningsguiden.pdf>
- XI, X. (2008). Methods of Test Validation. In N. H. Hornberger (Ed.), *Encyclopedia of Language and Education* (pp. 2316–2335). Springer US. Retrieved from <http://www.springerlink.com/content/x26628716822v82r/abstract/>
- Yildiz, L. M. (2011). *English VG1 level oral examinations*. (Master), The University of Oslo, Oslo. Retrieved from <http://urn.nb.no/URN:NBN:no-29277>
- Yle. (n.d.). Englanti syksy 2012 lyhyt oppimäärä. *abitreenit*. Retrieved November 5, 2012, from <http://abitreenit.yle.fi/yo-kokeet/Englanti/>
- Ylioppilastutkintolautakunta. (n.d.). The Finnish Matriculation Examination. *The Finnish Matriculation Examination*. Retrieved March 23, 2012, from <http://www.ylioppilastutkinto.fi/en/>

Zahedi, K., & Shamsaee, S. (2012). Viability of construct validity of the speaking modules of international language examinations (IELTS vs. TOEFL iBT): evidence from Iranian test-takers. *Educational Assessment, Evaluation and Accountability*, 24(3), 263–277. doi:10.1007/s11092-011-9137-z

Ørevik, S. (forthcoming 2012). *From “essay” to “personal text”: The role of genre in Norwegian EFL exam papers 1996-2011*. Forthcoming in *Acta Didactica Norge*.

Appendix 1

ENGLISH

WRITTEN PART

PITKÄ OPPIMÄÄRÄ
LÅNG LÄROKURS

13.3.2009



YLIOPPILASTUTKINTOLAUTAKUNTA
STUDENTEXAMENSNÄMNDEN

1 READING COMPREHENSION

- 1.1** Read texts 1.1a–1.1c and then answer questions 1–25. Choose the **best** alternative for each item and mark your answers **on the optical answer sheet in pencil**.

1.1a From Undercover to Between Covers

Joseph Weisberg looks about how you would expect a Brooklyn dad and schoolteacher to look, with a bald head, white-flecked beard and baggy leather jacket. So on a recent frigid night, when he ambled down a Park Slope street and stealthily passed off a plastic container from a gumball machine to a reporter, nobody noticed.

It was one of many examples of the spy tradecraft that Mr. Weisberg, 42, learned while training to be a case officer with the Central Intelligence Agency in the early 1990s. He no longer works there (or so he says), but he has used some of what he learned to write his latest novel, *An Ordinary Spy*.

The novel explores the moral complexity and psychological fallout of clandestine service, through a tight plot involving two case officers who meet after messed-up foreign assignments a few years apart. Told in restrained prose that reflects the emotional reserve of the characters, the book is more than a thriller. It is also a chronicle of the monotony of a spy's daily routine – not just the surveillance-detection routes and cryptic cables to headquarters, but also the staff meetings, petty rivalries between colleagues and idle chatter about pension plans. It's not quite "The Office" of espionage, but it's close.

"My goal was to write the most realistic spy novel that had ever been written," said Mr. Weisberg, who teaches history and English at a high school in Queens. To heighten authenticity, he edited parts of the book, inserting black bars that concealed the names of countries, the particulars of tradecraft and other details that might be classified information, if the story were true. Even the names of two people in the acknowledgments, whom the author thanks for having "trusted me with their story," are blacked out.

As a former CIA employee, Mr. Weisberg was required to submit his novel to the agency's Publications Review Board before he could even seek a book deal. During the five weeks that he waited to get the manuscript back from the board, he recalled over

35 lunch at a Brooklyn bistro, “I was convinced it was taking so long because they were taking out everything that was left.” In the end, he said, the board redacted only a little more than what he had already taken out. After he landed a publisher and went through rounds of edits, he resubmitted it – all told, a total of six times.

40 Anticipating the board’s rulings, Mr. Weisberg removed a lot of detail about basic spycraft. So it came as a surprise when the agency granted him permission to show a reporter how to perform a “dead drop,” a system of exchanging secret information without having to meet. He had folded explicit instructions into the aforementioned gumball machine container, providing directions
45 to the location of a signal, a chalk mark streaking across the side of a mailbox, and the actual drop, a note inside a crumpled chocolate bar wrapper and placed in a large flowerpot on a residential street.

The novel is narrated by Mark Ruttenberg, a young case officer
50 on his first assignment in an unspecified country. After he sleeps with a woman he had been trying to recruit as an informant, he is promptly fired. He soon gets to know another former officer, Bobby Goldstein, who had been stationed in the same country. The rest of the novel tells Bobby’s story, in which the exposing of a foreign
55 agent leads to consequences that Bobby is still struggling to live with.

Just before Mr. Weisberg was scheduled to leave on his first foreign assignment, he learned that his father’s cancer had become worse and took a leave of absence to help care for him. He returned
60 to a desk job to prepare for another foreign assignment, but instead left the agency altogether, although he chose to remain undercover. Friends said he had been telling them that he worked at the State Department.

When Mr. Weisberg first thought of joining the CIA, the Cold
65 War was still going on. Since graduating from Yale in 1987, he had taken some Russian classes, traveled around Eastern Europe and worked as a job counselor. Looking for something more meaningful, he remembered his early love for the spy novels of John le Carré.

70 Once he left the agency, Mr. Weisberg moved home in the mid-1990s to live with his mother in Chicago, where he wrote his first novel, about terrorists who blow up Wal-Marts. When he met and married his wife, Julie Rothwax, he started writing *An Ordinary Spy* and his money ran out. (His mother had been subsidizing him

75 until then.) So Mr. Weisberg took the teaching job. "It was the
only thing, other than writing, that interested me," he said. And
spying, of course.

Source: *The New York Times*, 2007

1.1b Attack of the Aliens

Tall and tan and fat and ugly, *achatina fulica* is not something
you'd want to behold on the sands of Ipanema. But you may not
80 have a choice. Since gaining a beachhead in Brazil 19 years ago,
this unlovely mollusk, better known as the giant African land snail,
has proved unstoppable. Imported secretly in 1988 as a cheap
substitute for escargot, it has caused widespread devastation,
blistering Latin America's biggest country like a strange pox.
85 Growing to the size of a man's fist and weighing one kilogram or
more, it lays up to 2,000 eggs a year and eats a tenth of its body
weight a day, devouring everything from lettuce to mouse
droppings to its own dead comrades. Worse, it can also carry rat
lungworm, a nasty parasite that burrows into the human brain.

90 Brazil's unwelcome snail is just one of a booming breed of
pests and pathogens that have broken free of their native habitats
around the world. Biologists somewhat quaintly call them exotic
species. The rest of the world knows them for what they are:
bioinvaders. They come in all guises, from trifling microbes to
towering trees, from the mosquito to the mongoose. What they
95 share is a tendency to stealth, spiriting around the planet on the
wings of migratory birds, nestled in the threads of clothing or
swimming in the human bloodstream. Scores of bioinvaders are
deliberately set loose by farmers or eggheads trying to outsmart
nature (importing snakes to chase rats), gardeners with exotic
100 tastes (knotweed) or entrepreneurs too clever by half (seeking the
perfect escargot). Bioinvaders are also ferocious competitors; free
from the predators of their homelands, they prosper on virgin
territory, monopolizing food supply and reproducing at a rate that
would make rabbits turn pale. Once lodged in a new land, the
105 intruding species may never go away, forcing public authorities to
battle them again and again with earthmoving equipment, fire and
poison – a job as futile as Sisyphus'.

110 There is nothing new or automatically harmful about wandering
wildlife. Without the millennial scrambling of life forms, human-

kind would neither eat nor prosper. More than 90 percent of food crops like wheat, corn and rice, and almost as many strains of livestock, are exotic species. But bioinvasion has taken a quantum leap in a borderless world where billions of people and tons of goods traverse the globe in a matter of hours, making a mockery of customs inspectors and quarantines. Indeed, the very forces that make the international economy flourish – trade, travel, transport and tourism – also make it vulnerable to invasive species.

Any pest, domestic or foreign, can be a nuisance, spoiling the flower patch or buzzing about the ears. But bioinvaders are especially dangerous. Some wipe out harvests, choke waterways and dry out the landscape, inviting wildfires. A deadly few microbes cause pandemics, like mad cow disease and AIDS. Even when they aren't an outright menace, exotic plants, animals and pathogens impoverish nature by crowding out a whole suite of homegrown species or creating mongrel hybrids through interbreeding.

Importing nature can be a blessing. A parasitic wasp from South America has helped millions of African farmers control the cassava mealybug, which ravages that continent's staple food, while Australia has successfully turned a killer virus from the Czech Republic against the omnipresent European rabbit. Often, though, nature bites back. The Indian mongoose was shipped to the West Indies to hunt rats, but ended up feasting on almost everything that crawled or croaked; a handful of ground nesting birds and up to a dozen amphibians and reptiles were driven to extinction.

Perhaps the only sure way to curb bioinvasion is to plug the gaping holes on international borders. If customs inspection in the United States is lax, in much of the rest of the world it's almost laughable. Only in 2005 did India get around to asking arriving passengers whether they were carrying any fruits, vegetables or plants – all major pathways for disease. But customs controls have their limits in the global economy. Thanks to tough import laws, Australia has developed one of the best bioinvasion defenses of any nation. Yet in the late nineties Canadian salmon farmers cried foul, charging Australia with unfair trade barriers. The World Trade Organization agreed, forcing Australia to open its market – a ruling that scientists fear could undermine other quarantine regulations.

And yet in a time when germs ride the wind and tide, even the most zealous border guards may be of little use. In most countries, exotic species are simply too established to eradicate. But that doesn't mean scientists should give up and grab the flamethrower.

The issue is not stopping bioinvasion, but understanding it. In the end, that means learning to live with the enemy.

Source: *Newsweek*, 2007

1.1c Home truths about telecom

Such is the social significance of mobile phones that when it comes to evaluating their use and planning new products and services, mobile operators cannot rely on the technology-driven, engineering mindset that has traditionally dominated the telecom industry. Most famously, industry leaders expected people to embrace videotelephony, which flopped, but failed to anticipate the success of text messaging. So they are turning to social scientists the better to understand how communications technologies are used.

Stefana Broadbent, an anthropologist for Switzerland's largest telecoms operator, has been looking at usage patterns associated with different communications technologies. She based her research on observation, interviews, surveys of users' homes and asking people to keep logbooks of their communications usage in several European countries. Some of their findings are quite unexpected.

People are in fact using different communications technologies in distinct and divergent ways. The fixed-line phone "is the collective channel, a shared organisational tool, with most calls made 'in public' because they are relevant to the other members of the household," she says. Mobile calls are for last-minute planning or to co-ordinate travel and meetings. Texting is for "intimacy, emotions and efficiency". Email is for administration and to exchange pictures, documents and music. Instant-messaging (IM) and voice-over-internet calls are "continuous channels", open in the background while people do other things.

Another finding is that despite the rise of free internet-calling services such as Skype, people seem to prefer typing. "The most fascinating discovery I've made this year is an increase in written channels," says Ms Broadbent. Her research in Switzerland and France found that even when people are given unlimited cheap or free calls, the number and length of calls does not increase significantly. This may be because there is only so much time you can spend talking; and when you are on the phone it is harder to do other things. Written channels such as email, text messaging

and IM, by contrast, are discreet and allow contact to be continuous during the day.

190 And although there is concern about work invading private life, the opposite actually seems to be true: private communications are invading the workplace. Workers expect to be plugged into their social networks while at work. Last year at a food-processing
195 factory near Geneva, the workers revolted when the director tried to ban mobile phones from the factory floor, and he was forced to relent. Their argument was that they wanted to be reachable during the day, just as people who sit at desks are.

Of course, improvements to mobile networks and the spread of third-generation networks mean that you no longer need to be at
200 your desk to get things done. But Ms Broadbent found that there is not, in fact, much appetite for working while on the move. After studying workers who spend more than half their time out of the office – salesmen, consultants, pilots, journalists and photographers – she found that they generally stick to gathering information while
205 on the move that they then work on when they get back to their desks. Hotel rooms and airports are not seen as an appropriate environment for substantive work and are mainly used for email.

Finally, Ms Broadbent found that migrants are the most advanced users of communications technology. A family of
210 immigrant workers from Kosovo living in Switzerland has installed a big computer screen in their living room, for example, and almost every morning they have breakfast with their grandmother back home, via a webcam. It is migrants, rather than geeks, who have emerged as the “most aggressive” adopters of new communications tools. Dispersed families with strong ties and limited
215 resources have taken to voice-over-internet services, IM and webcams, all of which are cheap or free. They also go online to get news or to download music from home. In the case of a Spanish family living in Switzerland, the daughter often does her homework
220 with her aunt but over a free Skype video-link, since the aunt lives in Spain.

Source: *The Economist Technology Quarterly*, 2007

1.2 Suomenkieliset koulut:

Lue seuraavat tekstit ja vastaa niiden pohjalta lyhyesti suomeksi kysymyksiin a–e. Kirjoita vastauksesi selvällä käsialalla kieli-kokeen vastauslomakkeen A-puolelle.

Svenska skolor:

Läs följande texter och ge sedan ett kort svar på svenska på frågorna a–e. Skriv svaren med tydlig handstil på sida A av svarsblanketten för språkproven.

Mussel farming in the UK is environmentally friendly, sustainable and very successful – so good news all round. Several different methods are used: often they are rope-grown, hanging like huge bunches of grapes on long ropes trailing in the water, while the French tend to favour the stake method, where they are attached to long poles standing upright in the coastal shallows.

People worry about eating shellfish, but you should feel confident about mussels; the water quality around the UK coastline is checked carefully and frequently, and the mussels are also inspected for bugs. Once harvested, they are purified in controlled tanks to rid them of any bacteria. Finally, anyone selling mussels has to display a health certificate for traceability.

Source: BBC Good Food, 2007

- a) Miten brittiläisten ja ranskalaisten simpukanviljelytavat eroavat toisistaan?
Hur skiljer sig de brittiska och franska sätten att odla musslor från varandra?
- b) Miten keräämisen jälkeen taataan, että simpukoiden syönti on turvallista? (Kaksi tapaa.)
Hur garanteras det efter skörden att det är tryggt att äta musslor? (Ange två saker.)

Sirs,

How interesting that the front page of your Canadian edition of *Time* [Oct. 1st] has the U.S. flag outstanding, out of all other flags, with the Canadian one being the least visible, on a piece of ice that belongs to Canada. There is no extent to America's "supremacy over the world" syndrome and the shameful lack of knowledge of your northern neighbor. How interesting that

(continued on page 11)

1.1a From Undercover to Between Covers

1. Why didn't Joseph Weisberg's actions draw any attention in the street?
 - A He tried to work skillfully
 - B He didn't do anything special
 - C He looked so ordinary
2. What did Weisberg do in the CIA in the early 1990s?
 - A He was taught how to write about spying
 - B He trained case officers
 - C He was coached for the spy business
3. What makes Weisberg's novel different from other thrillers?
 - A It tells about the private lives of spies
 - B It doesn't ignore the chores of espionage
 - C It concentrates on office work
4. What did Weisberg do to make his novel look more realistic?
 - A He pretended that it had been censored
 - B He wrote about the details of espionage
 - C He revealed some classified information
5. How did the CIA react to his novel?
 - A It edited it with a heavy hand
 - B It made a deal with Weisberg
 - C In no particular way
6. What was the dead drop Weisberg was allowed to show?
 - A Some candy wrapper
 - B Hiding info in plain sight
 - C Instructions how to find a certain mailbox
7. What was Bobby Goldstein's tragedy?
 - A He had been fired from the CIA
 - B He had given somebody up
 - C He had been revealed to be a spy
8. Why didn't Weisberg ever work for the CIA abroad?
 - A He was prevented by private affairs
 - B He was considered unreliable
 - C He was never given a chance

9. What inspired Weisberg to write spy novels?
 - A His travels in Eastern Europe
 - B His youthful interest
 - C His experience of terrorism
10. When did Weisberg start writing *An Ordinary Spy*?
 - A When he had no money left
 - B At the time of his marriage
 - C When he moved out from his mother's

1.1b Attack of the Aliens

11. What is said about the giant African land snail?
 - A It landed in Brazil accidentally
 - B It may be dangerous to man
 - C It eats mice and other animals
12. How do bioinvaders spread?
 - A They are illegally imported
 - B In many different ways
 - C They are carried by microbes
13. What makes bioinvaders dangerous?
 - A There's nothing to threaten them
 - B They kill domestic animals
 - C They can't be destroyed
14. What positive effect have bioinvaders had?
 - A They make the world a smaller place
 - B They provide food for people
 - C They enrich the environment
15. In what way may bioinvaders harm nature?
 - A By destroying some native species
 - B By causing the death of endangered animals
 - C By creating totally new diseases
16. What is said about importing nature?
 - A If not done carefully, it may be useless
 - B Parasites have spread in Africa
 - C Australia has had some success

17. Why is the prevention of bioinvasion problematic?
 - A It may break the laws of some countries
 - B It may violate the free flow of goods
 - C It may cause problems in customs inspections
18. What can man do about bioinvasion?
 - A Try to fight back with all possible means
 - B Try to make the best of the situation
 - C Try to regulate the trade of exotic species

1.1c Home truths about telecom

19. Why are mobile operators hiring social scientists?
 - A To have better communication with industry
 - B To understand better what people need
 - C To please the industrial leaders
20. How did Stefana Broadbent find out about people's usage patterns?
 - A She interviewed the operators
 - B She looked at how people use logbooks
 - C She had several research methods
21. What did Broadbent find out about the use of telecom?
 - A That mobile calls are the most popular
 - B That people are flexible in their usage
 - C That collective channels are becoming useless
22. What did Broadbent find surprising about free internet calls?
 - A They have not replaced written messages
 - B They have lost their attraction
 - C They are accessed by so few people
23. What is said about work and private life?
 - A Private affairs should not interfere with work
 - B Employees want to maintain social contacts on the job
 - C Workers should be available at any time
24. What do workers feel about working on the move?
 - A They think it saves valuable time
 - B They are not particularly fond of it
 - C They find it more strenuous
25. Why are immigrants keen on using telecom?
 - A It is affordable
 - B They want to show off
 - C It offers new challenges

America still considers itself the belly-button of the world, after all its unsuccessful attempts to run the world its way.

Mary Davis, Montreal

Source: *Time*, 2007

- c) Mistä kirjoittaja on alun perin tuhtunut, ja minkä yleisemmän väitteen hän esittää?
Vad var skribenten ursprungligen upprörd över, och vad påstår hon allmänt?

One of the most remarkable Chinese women to first settle in Britain is Song Ling Whang. At the beginning of the 20th century she made the long overland journey from China to Britain, on foot – a journey of over 6,000 miles. She walked with a group of other young people following the route of the trans-Siberian railway line to Europe, performing acrobatics and making paper flowers to earn their way. What is even more remarkable is that she made this journey on tiny claw-like feet that had previously been bound in the traditional Chinese manner.

Source: *BBC History Magazine*, 2007

- d) Miten Song Ling Whang löysi Eurooppaan, ja mitä kirjoittaja erityisesti ihmettelee?
Hur fann Song Ling Whang vägen till Europa, och vad undrar skribenten särskilt över?

As prey animals, horses in the wild need to be able to see as much as possible in order to spot predators early, so they can outrun them. The horse's eyes have therefore evolved over millions of years in order to give him the best chance of survival. Positioned near the top of the head, they allow him to see over the grass as he is grazing, so he can watch out for potential predators at all times. Because the eyes are also positioned on either side of the head, horses have two small blind spots: one just in front of the nose and another directly behind them. It's the reason we are taught never to approach a horse directly from behind, as he could be startled if you suddenly appear 'out of nowhere' into his line of sight.

Source: *Horse & Rider*, 2007

- e) Mitä positiivista ja mitä negatiivista hevosen silmien sijainnista seuraa?
Vad är positivt och vad är negativt med placeringen av hästens ögon?

2 GRAMMAR AND VOCABULARY

- 2.1** Read the text carefully and for each item choose the alternative that **best** fits the context. Mark your answers (26–55) **on the optical answer sheet in pencil**.

The Wait

Eisenhower was faced with a dreadful dilemma. On May 17, 1944, he had decided that D Day would 26 one of three days in June – the fifth, sixth or seventh. Meteorological studies had shown that two of the 27 weather requirements for the invasion could be 28 for Normandy on those days: a late-rising moon and, shortly after dawn, a low tide.

The paratroopers and glider-borne soldiers who would start the attack needed the moonlight. But their surprise attack 29 on darkness up to the time 30 arrived over the dropping zones. Thus their critical demand was for a late-rising moon.

The seaborne landings 31 to take place when the tide was low

26. A has to be
B was to be
C have to be
D were to be

27. A vital
B suiting
C fitting
D passable

28. A waited
B excepted
C expected
D accepted

29. A based
B required
C depended
D needed

30. A that
B which
C when
D they

31. A must
B had
C should
D was

enough to expose Rommel's beach obstacles. On this tide the timing 32 invasion would depend. And to complicate the meteorological calculations 33, follow-up troops landing much later in the day would also need a low tide – and it had to come before darkness 34 in.

These two critical factors of moonlight and tide shackled Eisenhower. Tide 35 reduced the number of days in 36 one month for the attack to six, and three of those were moonless.

But 37. There were many other considerations he had to take into 38. First, all the services wanted long hours of daylight and good visibility – to identify the beaches; for the naval and air forces to spot their targets; and to reduce the hazard of collision when five 39 ships began maneuvering almost side by side in the bay of 40 Seine. Second, a

32. A the whole
B of whole
C of the whole
D whole of the
33. A farther
B further
C furthermore
D moreover
34. A fell
B came
C set
D became
35. A just
B even
C only
D alone
36. A some
B any
C a
D every
37. A enough with that
B on top of everything
C that was not all
D notwithstanding
38. A notice
B mind
C observation
D account
39. A thousand
B thousands
C thousands of
D of thousands
40. A a
B an
C the
D –

calm sea was required. 41 from the havoc a rough sea might cause to the fleet, seasickness could leave the troops helpless long before they even set 42 on the beaches. Third, low winds, blowing inshore, were needed to 43 the beaches of smoke 44 targets would not be obscured. And finally, the Allies required three more quiet days after D Day 45 the quick build-up of men and supplies.

46 at Supreme Headquarters expected perfect conditions on D Day, least of all Eisenhower. He had schooled himself, in countless dry runs with his meteorological 47, to recognize and weigh all the factors which would give him the bare minimum conditions acceptable for the attack. But, 48 his meteorologist, the 49 were about ten to one against Normandy having weather on any one day in June

41. A In addition
B Besides
C Apart
D Considering
42. A eye
B foot
C leg
D hand
43. A wipe
B clear
C clean
D lift
44. A in case
B because
C unless
D so that
45. A would facilitate
B to facilitate
C facilitating
D facilitated
46. A Anybody
B Nobody
C Everybody
D All
47. A staff
B stuff
C persons
D personage
48. A after
B as far as
C in accord with
D according to
49. A choices
B options
C chances
D changes

which would 50 even the minimal requirements.

Of the three possible days for the invasion he had chosen the fifth so that if 51 was a postponement he could start the attack on the sixth. But if he ordered the landings for the sixth and then had to cancel again, the problem of refueling the returning convoys might prevent him from attacking on the seventh. There would then be two alternatives. He could postpone D Day until 52 period when the tides were right, June 19. But if he 53 that, the airborne armies would be forced to attack in darkness – June 19 was moonless. 54 alternative was to wait until July. And that 55 postponement, as he was later to recall, “was too bitter to contemplate.”

50. A answer
B fill
C meet
D respond

51. A it
B then
C there
D –

52. A next
B the next
C following
D the following

53. A would
B should
C did
D had to

54. A An other
B Other
C The other
D One

55. A long
B a long
C long a
D the long

Source: Cornelius Ryan, 1999
The Longest Day June 6, 1944

- 2.2** Fill in the blanks using suggestions when given. Write your answers in the given order **on side B** of the answer sheet. Write each answer on a separate line. Please write **clearly**.

The Recipe for Happy Cats

You may be surprised to learn that 45 per cent of UK cat-owning households have more than one cat – and if 1 is one of them, you'll know how stressful mealtimes can be for them – and for you! If your cats are at different life stages, or one is overweight and one is thin, it can be pretty tricky making sure each individual pet 2 the most nutritious food 3 its age and health.

1. sinun/ditt

2. feed

3. prepositio/
preposition

Food is not the only area that can cause problems for the modern feline in a multi-cat household. Homes have become much less cat-friendly of late, with the fashion for minimalism meaning that 4 harder for cats to find cosy corners or good vantage points from which 5 their surroundings.

4.

5. observe

Jan Cisek says, 'A happy cat really does equal a happy owner and a happy

home.' For example, owners should place several food stations around the house 6 their cats to follow natural grazing patterns. Cat owners should also think about swapping leather furniture for softer fabrics to instill a sense of 7 and relaxation in their pets.

6. rohkaisemaan /
för att uppmuntra

7. turvallisuus/trygghet

Here are some tips to transform your household into a feline home:

- Don't be too tidy. Cats actually like a bit of clutter as they love to hide and explore in nooks and crannies.
- Don't place 8 food and water bowls 9.
- Cats like these elements to be separate – this will also make them want to drink more from their bowls and less from the tap!
- Make sure you create some power positions for your cats. Places that are high up, from where they can observe their world, make for a truly 10 animal.

8. kissojesi / dina katters

9. vierekkäin /
bredvid varandra

10. tyytyväinen/belåtet

Source: *BBC Good Food*, 2007

3 PRODUCTION

*Write a composition of between 150 and 250 words on one of the following topics. Please write **clearly** on the notebook paper (konseptipaperi/konceptpapper) provided. Follow the guidance. Count the number of words in your essay and write it at the end.*

1. What am I?

All of us are the sum of our genes and the environment around us. Do genes or the environment influence you more strongly? How does this show in you personally?

2. Living in the past, living in the future

Large numbers of youth dress up for role playing on weekends. What do you think are the reasons for this interest? Which do you prefer, the past (for example, medieval times) or the future (for example, science fiction)?

3. A speech

Your family has asked you to give a speech at your Grandmother's or Grandfather's 80th birthday. Write this speech.

4. Why extreme? How extreme?

Why do some people risk their lives by doing extreme sports? Would you take up, or have you taken up, an extreme sport? Why or why not? Should some extreme sports be banned?

KOKEEN PISTEITYS / POÄNGSÄTTNINGEN AV PROVET

Tehtävä	Osioiden määrä	Pisteitys	Paino-kerroin*	Enint.	Arvostelu-lomakkeen sarake
Uppgift	Antal deluppgifter	Poäng-sättning	Koefficient*	Max.	Kolumn på bedömnings-blanketten
1.1a–c	25 x	1/0 p.	x 2	50 p.	1
1.2	5 x	2–0 p.	x 2	20 p.	2
2.1	30 x	1/0 p.	x 1	30 p.	3
2.2	10 x	1–0 p.	x 1	10 p.	4
3				99 p.	7
Yht./Tot.				209 p.	

*Painotus tapahtuu lautakunnassa.
Viktningen görs av nämnden.

Appendix 2

ENGLISH

WRITTEN PART

PITKÄ OPPIMÄÄRÄ
LÅNG LÄROKURS

17.3.2010



YLIOPILASTUTKINTOLAUTAKUNTA
STUDENTEXAMENSNÄMNDEN

1 READING COMPREHENSION

1.1 Read texts 1.1a–1.1d and then answer questions 1–25. Choose the **best** alternative for each item and mark your answers **on the optical answer sheet in pencil**.

1.1a A scientific detective story

As collections manager at London's Natural History Museum, Max Barclay has traveled the world in search of rare and previously undiscovered insects. So when his 5-year-old son took a break from a picnic lunch last March in the museum's garden and returned with an insect in his hand, Barclay could not have guessed that his question –
5 "Daddy, what's this?" – would lead to a global detective hunt that has so far baffled Barclay and the world's other entomologists.

Despite working with an insect collection of more than 28 million specimens, Barclay and his colleagues have been unable to identify
10 the almond-shaped critter, about the size of a grain of rice, which has in the past year made itself at home in the sycamore trees on the 19th century museum's grounds in central London. "My field work has taken me all over the world – to Thailand, Bolivia, Peru. So I was surprised to be confronted by an unidentifiable species while having
15 a sandwich in the museum's garden," Barclay says.

Within three months of the discovery, the insect had become the most common species in the garden and was spotted in other central London parks, sending Barclay on a worldwide hunt to identify it. Correspondence with colleagues around Europe led Barclay
20 to discover that the insect, which resembles the common North American box elder bug, is actually most closely related to *Arocatus roeselii*, a relatively rare species of seed eaters usually found in central Europe. But those bugs are associated with alder trees rather than sycamores. An insect specimen found in Nice, France, which
25 is now in the collection at the National Museum in Prague, turned out to be the same as the mysterious London bug. But that specimen had been misidentified as *Arocatus roeselii*. "There are two possible explanations," says Barclay. "One is that the bug is *roeselii*, and by switching to feed on the sycamores, it has suddenly become more
30 abundant, successful and invasive. The other is that the insect in our grounds may not be *roeselii* at all."

In recent years, several foreign insects, spiders and beetles have been discovered in Britain, a trend many attribute to the ability of

35 such species to survive winters warmed by climate change. In 2005
Edinburgh Zoo issued a public notice after several panicked Scots
reported seeing a spider called a false widow, which has a disconcerting
behavior of rushing toward people who approach it. The spider turned
out to be quite common – in the Canary Islands.

40 Barclay is not convinced that climate change is responsible for
Britain's new inhabitants. European integration may be the cause. "It's
very difficult to judge," Barclay says, "because the period of time we
have seen global warming potentially influencing the insect fauna
is almost exactly the same period of time since the European Union
opened up its trade barriers between member states. So in the past
45 decade and a half, we've been importing a lot more from Italy and
Spain and Southern France, and we've had this climatic change – so
we have two potential causes." Whatever the reason for the appearance
of the new bug in London, Barclay says its spread is harmless. But
he concedes, "It does show what's possible if more damaging species
50 invade."

The struggle to identify the museum insect displays not only the
mystery of nature but also the fickleness that surrounds the science
of taxonomy. Figuring out which insects are which can be fiendishly
difficult; some scientists estimate that we have managed to identify
55 only 10% of the insect world so far. The rest, like Barclay's almond-
shaped mystery bug, are perfectly happy to crawl along without any
christening or approval from their gargantuan neighbors. But that won't
stop scientists like Barclay from trying to give his new chums a proper
name – that is to say, a Latin one. For Barclay, the question asked by
60 his son last March amounts to a calling he still feels compelled to
answer.

Source: *Time*, July 17, 2008

1.1b The Dumbest Generation? Don't Be Dumb.

Really, don't we all know by now that finding examples of teens' and
twentysomethings' ignorance is like shooting fish in a barrel? If you
want to exercise your eye-rolling or hand-wringing muscles, take your
65 pick. Two thirds of high-school seniors in 2006 couldn't explain an old
photo of a sign over a theater door reading COLORED ENTRANCE.
In 2001, 52 percent identified Germany, Japan or Italy, not the Soviet
Union, as America's World War II ally.

70 Like professors shocked to encounter students who respond with a
blank-eyed "huh?" to casual mentions of Pearl Harbor, and like parents

appalled that their darling doesn't know Chaucer from Chopin, Mark Bauerlein sees in such ignorance an intellectual, economic and civic disaster in the making. In his provocative new book "The Dumbest Generation: How the Digital Age Stupefies Young Americans and Jeopardizes Our Future (Or, Don't Trust Anyone Under 30)," the professor of English offers the usual indicators, grand and slight. He gives as evidence a decline in adult literacy and a rise in geographic cluelessness.

He is a little late to this party, of course. The old have been wringing their hands about the young's cultural wastelands and ignorance of history at least since admirers of Sophocles and Aeschylus bemoaned the popularity of Aristophanes as leading to the end of Greek civilization as they knew it. Victorian scholars considered Dickens, that plot-loving, sentimental favorite, a lightweight compared with other authors of the time. Civilization and culture, high and low, survived it all. Can it survive a generation's ignorance of history? For those born from 1980 to 1997, Bauerlein lamented to us, "there is no memory of the past," just like when the Khmer Rouge said "this is day zero." Historical memory is essential to a free people. Fair enough, but we suspect that if young people don't know the import of old COLORED ENTRANCE signs it reflects not stupidity but a failure of the school system and of society (which is run by grown-ups) to require them to know it. Drawing on our own historical memory also compels us to note that philosopher George Santayana, too, despaired of a generation's historical ignorance, warning that "those who cannot remember the past are condemned to repeat it." That was in 1905.

A more fundamental problem is what Bauerlein has in mind by "dumbest." If it means "holding the least knowledge," then he has a case. Gen(eration) Y cares less about knowing information than knowing where to find it. A few keystrokes would easily give you an answer to any information you need. But if dumb means lacking such fundamental cognitive capacities as the ability to think critically and logically, to analyze an argument, to learn and remember, to see analogies, to distinguish fact from opinion ... well, here Bauerlein is on shakier ground.

First, IQ scores in every country that measures them, including the United States, have been rising since the 1930s. Since the tests measure not knowledge but pure thinking capacity – what cognitive scientists call fluid intelligence, in that it can be applied to problems in any domain – then Gen Y's ignorance of facts (or of facts that

older people think are important) reflects not dumbness but choice. And who's to say they are dumb because fewer of them than of their grandparents' generation care who wrote the oratorio "Messiah"? Similarly, we suspect that the decline in the percentage of college freshmen who say it's important to keep up with political affairs, from 60 percent in 1966 to 36 percent in 2005, reflects at least in part the fact that in 1966 politics in the US determined whether you were going to get drafted and shipped to Vietnam. The apathy of 2005 is more a reflection of the world outside Gen-Yers' heads than inside. Alienation is not dumbness.

Bauerlein is not the first scholar to pin the blame for a younger generation's intellectual shortcomings on new technology (television, anyone?), in this case judging "the digital age." But there is no empirical evidence that being immersed in instant messaging, texting, iPods, videogames and all things online impairs thinking ability. The jury is still out on whether these technologies are positive or negative for cognition. But they're definitely changing how people's brains process information.

Source: *Newsweek*, June 2, 2008

1.1c Prickly charmers

Aristotle believed hedgehogs could predict the weather. Arthur Schopenhauer, philosopher, used them to illustrate the challenges of human intimacy. Beatrix Potter's character Mrs Tiggy-Winkle has enchanted children and adults alike since 1905, while Sonic, her modern-day counterpart, is one of the world's best-known video game characters.

In this engaging memoir *Charming Mammal* Hugh Warwick draws on 20 years of knowledge to explain why hedgehogs have gained such iconic status. He sprinkles his book with facts: hedgehogs snooze through the winter in an aptly named hibernaculum; they have up to 7,000 spines; their fleas are species specific; North America has no native hedgehogs although thousands of imported ones are kept as pets.

Mr Warwick also introduces several characters whose obsession with hedgehogs is all-encompassing. Most memorable is Zug Standing Bear, a former war-crimes investigator, one-time bodyguard to Gerald Ford and now owner of Buttercup, whose athletic prowess won her a gold medal at the International Hedgehog Olympic Games in Denver, Colorado. As they say, it takes all sorts.

Source: *The Economist*, December 30, 2008

1.1d Call of the Wild. The Howling Miller

Anyone familiar with the films of Ingmar Bergman or the writing of Halldor Laxness knows that grim hilarity is the stock in trade of many
150 a 20th-century Nordic artist. Maybe a little black humor enables a person to better endure the harsh climate and rugged terrain.

A case could certainly be made after reading the Finnish writer Arto Paasilinna's novel, which takes place in post-World War II Lapland. The protagonist, Gunnar Huttunen, arrives one day at a
155 sleepy, woodsy canton in the province and quickly sets up a shingling service run out of a decrepit mill. (Huttunen's scruffy mien mirrors Lapland's real-life heraldic symbol: a bearded, shirtless man holding a club across his shoulders.) His provenance is under suspicion, and he has a tendency to howl during his periodic "attacks of depression,"
160 but he's remarkably handy and industrious. "He's crazy, but he splits good tiles and they're reasonably priced," the local farmers conclude. (The author himself bears a workaday resemblance to Huttunen: his publishers identify him as "by turns a woodcutter, agricultural laborer, journalist and poet" – and, no slouch himself, he's written more than
165 20 novels.)

The narrative, with its episodic chapters and fablelike tone, follows Huttunen's interactions with an amused eye. Prescribed a potent drug to cure his howling, Huttunen impetuously downs a handful of pills and goes on a midnight rampage through his neighbors' properties.
170 He is shipped off to a mental hospital in the nearby city of Oulu and declared a manic-depressive.

Paasilinna describes the frenetic inner workings of his characters' minds with an expert touch. In one finely wrought passage, Huttunen's female friend, Sanelma, imagines having Huttunen's baby: "He'd be
175 at least three foot at birth. No one would dare breast-feed him, this deranged cherub fathered by a lunatic. He wouldn't burble away like an ordinary newborn, but would howl like his father. Or whine at least. Normal children's clothes wouldn't fit him; she'd have to sew him sailor trousers to wear in his cot. He'd grow a beard by the time
180 he was 5." It is Paasilinna's gift in this gem of a novel to wring humor from the most desperate of circumstances.

Source: *The New York Times*, December 28, 2008

1.2 Suomenkieliset koulut:

Lue seuraavat tekstit ja vastaa niiden pohjalta lyhyesti suomeksi kysymyksiin a–e. Kirjoita vastauksesi selvällä käsialalla kielikokeen vastauslomakkeen A-puolelle.

Svenska skolor:

Läs följande texter och ge sedan ett kort svar på svenska på frågorna a–e. Skriv svaren med tydlig handstil på sida A av svarsblanketten för språkproven.

Dark Horse

The future of biofuel may lie in one word: algae. The tiny primitive plants can produce a lot of oil in a little space. Solazyme in South San Francisco leases the vast fermentors used in industrial processes to grow specially tailored strains of algae in the dark. The algae swim in a bath of sugar water that delivers the sustenance they need to produce copious quantities of oil. The company can select different types of algae and alter the conditions they grow in to produce various oils that can be turned into diesel fuel, plastics or other products. "Some strains come from a swamp in Africa; some strains come from a high-altitude snowfield in South America," says Harrison Dillon, a microbiologist and co-founder.

The firm has already produced thousands of gallons of oil, though not at "commercial economics," as Dillon puts it. Notably, the company car, a diesel Jeep, runs on the stuff. And the U.S. Military and the American Society for Testing and Materials have certified the diesel for a broad range of vehicles; the company plans to open its first dedicated production facility later this year. Dillon says refining production and scaling up processes will allow Solazyme to make algal oil a feedstock for diesel fuel that will cost only \$2 per gallon. Because sunlight is not needed, improvements are easier to come by, he maintains. Solazyme may be the algae company furthest along in producing different oils at commercial volumes, which could potentially be used in everything from jets to kitchens.

Source: *Scientific American Earth* 3.0, 2008

- a) Mitä Solazyme valmistaa levistä ja miten?
Vad tillverkar Solazyme av alger, och hur?
- b) Miten pitkälle tuotteen kehittälyssä on päästy?
Hur långt har man kommit i produktutvecklingen?

Vertical farming

Why not grow grains, vegetables and fruits right where the expanding crowds of consumers are: in the middle of a city, inside a tall glass building? Poultry and pork could be reared there, too. A vertical farm would drastically reduce the fossil-fuel use and emissions associated with farm machinery and trucking, as well as the spread of fertilizer and its runoff. Crops could grow and be harvested year-round instead of at the end of one season, multiplying annual yield by at least four times. Urban agriculture could also convert municipal wastewater into irrigation water, reducing a city's refuse problem. And consumers would get the freshest food possible, without pesticides.

Sceptics worry that indoor crops would have to be genetically modified to thrive, but Dickson Despommier of Columbia University, who has championed the movement, says such alteration is unnecessary, "we can match their growth characteristics with temperature and humidity conditions and nutrition profiles."

Although real estate in cities would seem too pricey for farming, Despommier's graduate students surveyed New York City and found many abandoned buildings and lots where high-rise farms could sprout. "You can do this on the rooftops of hospitals and schools," which would use the food in their own kitchens. "You can do this along the periphery but still within city limits. You can do this on open stretches of air force bases and airports and city islands."

Source: *Scientific American Earth* 3.0, 2008

- c) Mitä tekstissä ehdotetaan, ja millä eri tavoin tämä hyödyttäisi nimenomaan kaupunkilaisia?
Vad föreslår man i texten, och på vilka olika sätt skulle detta vara till nytta i synnerhet för stadsbor?
- d) Miksi ehdotus herättää epäilyksiä?
Varför väcker förslaget tvivel?
- e) Mitä Despommier teetti opiskelijoillaan?
Vad lät Despommier sina studenter göra?

- 1.1a**
1. What is Max Barclay's job?
 - A He travels around the world to see insect collections
 - B He looks for items for his museum
 - C He is the main director of the Natural History Museum
 2. What did Barclay's son do?
 - A He ran away from the museum garden
 - B He dismayed the world's entomologists
 - C He found something interesting
 3. What do we learn about the unknown insect?
 - A Its size, shape and habitat
 - B Where it's from and what it eats
 - C That it had lived in London for years
 4. Why did Barclay contact colleagues worldwide?
 - A He was worried about the bug spreading fast
 - B The insect had caused much harm in London parks
 - C He wanted to know if anyone knew the insect
 5. Where can an identical bug be seen?
 - A In Nice, France
 - B In Prague, the Czech Republic
 - C In the National Museum in London
 6. What created a problem for the scientists?
 - A That the insect lived in another tree than the sycamore
 - B That the insect defied final identification
 - C That the insect was from the American continent
 7. Why had some Scots panicked?
 - A They had spotted a poisonous spider in Edinburgh Zoo
 - B They had been scared by a threatening spider
 - C They feared they had killed a rare spider
 8. What does Barclay think are the possible reasons for new species in Britain?
 - A Global warming and traveling
 - B European Union inhabitants and climatic change
 - C Freer trade and warmer climate

9. What challenge do entomologists still face?
A Few insects have other than Latin names
B Only a small minority of insects are known
C That many insects will soon die out
- 1.1b** 10. What is said about American high school students in the first paragraph?
A They irritate their elders by their bad manners
B Many don't have a very solid knowledge of history
C Many don't have any interest in schoolwork
11. What is new in Prof. Bauerlein's ideas?
A He sees the young differently
B Nothing really, they are age-old
C He concentrates on American youth
12. What does the writer think of young people's poor knowledge of history?
A Young Americans don't really care about history
B Society should take more responsibility
C There should be more history in schools
13. Why aren't the young interested in knowing facts?
A They are not used to holding on to knowledge
B They know that facts are easily available
C They know that facts change fast
14. What change has taken place in many countries over the past few decades?
A People's thinking abilities have improved
B People's educational level has risen
C The use of IQ tests has become more common
15. Why was it important for young American males to follow politics in 1966?
A There was an imminent threat of war
B The 60s was a very turbulent political decade
C They may have faced sudden life changes
16. Why are today's young Americans less interested in politics?
A Current politics doesn't relate to them
B There's nothing dramatic going on
C They are so apathetic

17. What is wrong with Bauerlein's accusation?
 A The bad influence of new technology hasn't been proved
 B He doesn't seem to know the latest technology
 C He gives false evidence for his arguments
- 1.1c** 18. What do Mrs Tiggy-Winkle and Sonic have in common?
 A They are both fictive figures
 B They are both loved by children
 C They both live in the woods
19. What makes hedgehogs special for many Americans?
 A They have different kinds of bugs
 B They sleep through the winter
 C They afford pleasure to many
20. Why is Buttercup famous?
 A She worked for Gerald Ford
 B She excelled in sports
 C She helped to solve war crimes
- 1.1d** 21. What is said about the Nordic sense of humor?
 A It has been influenced by weather and nature
 B It is hard for other nations to understand
 C It is hilarious and crazy
22. What do Huttunen and Lapland's heraldic symbol have in common?
 A They are both old-fashioned
 B They look like each other
 C They are both based on real-life figures
23. What is the local farmers' attitude towards Huttunen?
 A They appreciate his skills
 B They find him untrustworthy
 C They think he is reasonable
24. Why is Huttunen taken to a mental hospital?
 A He had tried to harm himself
 B He had taken strong drugs
 C He had attacked his neighbors
25. What is said about Huttunen's child?
 A He is much like his father
 B He does not exist
 C He is big and strong

2 GRAMMAR AND VOCABULARY

- 2.1** Read texts 2.1a and 2.1b carefully and for each item choose the alternative that **best** fits the context. Mark your answers (26–55) on the optical answer sheet in pencil.

2.1a Saving Private Ryan

I have discovered a fundamental law of teaching: it is almost impossible to print off enough copies of letters for pupils

_____ 26 _____ home to their parents or carers.

In fact, at a conservative estimate, you need to print 8,000 copies of each letter for each pupil. That way, perhaps one or two copies would _____ 27 _____ it home.

We had to send letters out to the Year 9s (_____ 28 _____ thirteen to fourteen) for their parents to give permission for us to show _____ 29 _____ the opening sequence from *Saving Private Ryan* in preparation for the school visit to Normandy. After several weeks of continuous reminders we eventually got the necessary slips in, and the showings commenced.

It is quite incredible how different classes from the same year can be. I ended up having to show the film to three out of the four groups. The first lot watched with transfixed interest, and shrieked _____ 30 _____ excitement when the US Rangers' sharpshooter aimed his rifle

26. A take
B to take
C taking
D taken

27. A do
B find
C make
D work

28. A age
B ages
C aged
D aging

29. A to them
B for them
C them
D those

30. A for
B in
C of
D with

and killed one of the enemy machine-gunners. The second class watched with equal fascination but sat in dumbfounded silence when we finished before gradually breaking into a moving and touching reflection on what they had seen. This was 31 the fact that I had let the film run to the point where the priest arrives to tell Private Ryan's mother that his three brothers 32 killed.

The third class were giggly to begin with, but the giggles soon stopped. I was sitting at the back and I hadn't noticed what effect it was having 33 I walked to the front just as the priest's car drew up. I was confronted with a sea of tear-sodden faces, which took me almost to the point of no return too. I packed them off five minutes early to go and ruminate in private.

Actually, they were fine about it, but I wondered – should we show them this sort of thing? And the answer has to be yes, because the world they live in is a product of that age, and we are not immune to it either. One of them told me the following day that 34 18-year-old brother is serving in Iraq. She only sees him twice a year and lives in continual fear of 35 might happen to him.

31. A apart
B despite
C in spite
D regardless

32. A are
B were
C have been
D had been

33. A after
B before
C until
D when

34. A his
B her
C their
D a

35. A something
B that
C what
D which

Saving Private Ryan's story might be fiction and shameless emotional manipulation, but the setting and circumstances aren't – thousands of parents received those horrible letters

36 6 June 1944. Omaha Beach was bloody for a reason and with our trip looming I knew the pupils would appreciate 37 the enormity of the undertaking because now they have an inkling of the sacrifices made then, and still being made by families today.

Source: *BBC History magazine*, October, 2008

2.1b Stronger than steel

For the opening race of the Eastern U.S. collegiate cycling season, Nick Frey, 38 the under-23 national time-trial champion, had left his \$13,000 carbon-fiber team bike at home. What he brought 39 was an even more recent model: a racing bike he and friends had made from bamboo.

Pioneers beyond the bike world are discovering that bamboo 40 be the most useful raw material ever to be overlooked. 41 a common building material in many tropical countries, it's considered "the poor man's timber," and in the West it's 42 decorative. As the

36. A after
B at
C since
D in
37. A much
B plenty
C more
D both
38. A newly
B currently
C constantly
D lately
39. A instead
B in case
C in place
D on the spot
40. A can
B may
C must
D will
41. A Yet
B But
C Although
D Even if
42. A primary
B mainly
C foremost
D barely

world 43 green, however, bamboo's essential qualities are beginning to win converts.

Despite its lowly reputation, bamboo may be the strongest 44 on the planet. It has greater tensile strength (or resistance to being pulled 45) than steel, and it withstands compression better than concrete. Both qualities are essential to keeping the plant, which grows to nearly 60 meters but is only as wide at the base as the very top, from falling over. Bamboo's strength is in its ability to bend, and that's 46 miracle.

The Western mind is also opening to bamboo's environmental qualities. In both temperate and tropical climates, it grows as quickly and abundantly 47 a weed. Though most often used as an alternative to timber, bamboo is technically a type of grass. In fact, it grows faster than 48 other grass – in some conditions well over a meter a day. It also 49 considerably more oxygen from carbon dioxide than trees and more effectively binds 50 to prevent erosion.

Bamboo is still a niche material in the United States, but it's catching on. Smith Fong first cracked the markets in 1989 with flooring and, later, a laminated bamboo sheet called "plyboo."

- 43. A is
B turns
C comes
D goes
- 44. A staff
B stuff
C matter
D constituent
- 45. A apart
B off
C out
D up
- 46. A a
B an
C the
D –
- 47. A like
B as
C than
D so
- 48. A one
B any
C no
D some
- 49. A demands
B produces
C needs
D requires
- 50. A earth
B ground
C soil
D land

Bike designer Craig Calfee says it's only a matter of time before bamboo sweeps the bike-racing world, and he 51 know. Calfee pioneered the use of carbon fiber, the gold standard for elite bicycle frames since the 52. Now he says bamboo may be 53, offering a better combination of stiffness (for power efficiency) and compliance (for vibration dampening).

Frey and three engineering students are picking up Calfee's torch, 54 to refine the design and reduce the price. Riding a bamboo bike, says Frey, "is like 55 comfortable loafers and having the efficiency of track spikes." It's a technological wonder – that nature already built.

Source: *Newsweek*, Oct. 2008

51. A may
B can
C ought
D should
52. A 1990
B 1990s
C 19th century
D 90 years
53. A super
B superb
C superior
D surplus
54. A aiming
B meaning
C going
D thinking
55. A to wear
B to use
C wearing
D using

- 2.2** Fill in the blanks using suggestions when given. Write your answers in the given order **on side B** of the answer sheet. Write each answer on a separate line. Please write **clearly**.

A Letter to My Son On Election Night

What does Barack Obama's election mean to you? It means many things. When you are older, we will talk about how African-

American children, like their parents and grandparents, have struggled to overcome the feeling that no matter how hard they study and work and try, there are barriers – 1 visible, others hidden but still there – that block their way. The feeling that we can rise, but only so far. I did not want you to grow up believing that bitter remnants of the past could hold power over your future. I wanted to be able to tell you that it wasn't true – that you could be 2 you wanted to be. But I couldn't quite believe it myself. Now I do.

1. toiset/några

2. mitä tahansa /
vad som helst

With Obama's election, I can mean it when I tell you that the world is available to you. No election can wipe away racism, and bigotry will show itself to you in ways subtle and not. But it is easier today than it was yesterday to see that racism, 3 a barrier, is now more like a hurdle. Barack Obama faced hurdles but succeeded, and you can, too. You are only four months old, but already I dream of what a great rocket scientist you'd be.

3. joka ennen oli /
som var tidigare

If you do become president, it won't be just because you 4 the votes of people who look 5 you. This election was such a triumph because it tested our nation's fundamental promise of equality – and we as a nation passed. Black Americans did not elect Obama.

4. win

5.

Americans did. 6 every African-American in the country had voted for him, it would not have been enough. He will enter 7 White House with the support and good wishes of millions of people of all races, colors and creeds.

Yet this hope of greater expectations for ourselves comes with greater expectations on ourselves. We can no longer look to blame outside forces 8 our failures. If you want to be like Obama, you have to work and work hard. He 9 not simply dream of being president; he summoned the discipline to get there.

I speak from experience. In eighth grade, I made a contract with myself: I was going to Yale. I worked hard in college and graduated. Along the way, there were 10 of people who doubted me and said unkind things, and not all of them were white. Don't let anyone mess with your mind or steer you from your goals. All of us have to defy stereotypes. Your dad is white, which means technically you are biracial – but that's black in America. If anyone doubts your "authenticity," you can tell them that you want to be as black as the man behind the desk in the Oval Office.

Source: *Newsweek*, November 17, 2008

6. Vaikka/Fastän

7.

8.

9.

10. paljon/många

3 PRODUCTION

*Write a composition of between 150 and 250 words on one of the following topics. Please write **clearly** on the notebook paper (konseptipaperi/konceptpapper) provided. Follow the guidance. Count the number of words in your essay and write it at the end.*

1. Dear Fellow Europeans

You will be participating in European Youth Week and will give a speech on a topic you find important to all young Europeans. Write this speech.

2. No way!

You live in England and there are plans to build a new motorway which will pass very close to your home. Write a letter to be published in a local paper giving your opinion about the plan.

3. Can one person change the world for the better?

Write about one person living today who you believe will do much good for the world. Defend your view.

4. James Bond – a hero of the past?

A film magazine has invited its readers to take part in a writing competition under the title above. The winning entry will be published. Write your article about James Bond for the magazine.

KOKEEN PISTEITYS / POÄNGSÄTTNINGEN AV PROVET

Tehtävä	Osioiden määrä	Pisteitys	Painokerroin*	Enint.	Arvostelulomakkeen sarake
Uppgift	Antal deluppgifter	Poängsättning	Koefficient*	Max.	Kolumn på bedömningsblanketten
1.1a–d	25 x	1/0 p.	x 2	50 p.	1
1.2	5 x	2–0 p.	x 2	20 p.	2
2.1	30 x	1/0 p.	x 1	30 p.	3
2.2	10 x	1–0 p.	x 1	10 p.	4
3				99 p.	7
Yht./Tot.				209 p.	

* Painotus tapahtuu lautakunnassa.
Viktningen görs av nämnden.

Appendix 3

ENGLISH

WRITTEN PART

PITKÄ OPPIMÄÄRÄ
LÅNG LÄROKURS

18.3.2011



YLIOPPILASTUTKINTOLAUTAKUNTA
STUDENTEXAMENSNÄMNDEN

1 READING COMPREHENSION

- 1.1** Read texts 1.1a–1.1c and then answer questions 1–25. Choose the **best** alternative for each item and mark your answers **on the optical answer sheet in pencil**.

1.1a Soaps, sex and sociology

The glamorised world portrayed on the nightly *telenovelas* (soap operas) on Brazilian television is, superficially at least, about as representative of the country as a whole as Marie Antoinette and her shepherdesses were of 1780s France. But they are all about aspiration.

- 5 About 40m people watch the mid-evening *novela* from Globo, the leading network. The action often takes place in Rio de Janeiro, where Globo is based, among families which are smaller, whiter and richer than average. New research suggests that by selling this version of the country to itself, Globo has boosted two important social trends.

- 10 The soaps blossomed under Brazil's military regime of 1964–85. The generals subsidised sales of television sets to build a sense of nationhood in a large and then largely illiterate country. National news was meant to do the job, but the soaps got the audience. Their scriptwriters and directors saw them as a tool with which to reach
15 the masses. Their plots often tilt in a progressive direction: AIDS is discussed, condoms are promoted and social mobility exemplified.

- How much impact do the soaps have on real life? As recounted in papers from the Inter-American Development Bank, researchers tracked Globo's expansion across the country and compared this to
20 data on fertility and divorce.

- The results are most striking for the total fertility rate, which dropped from 6.3 children per woman in 1960 to 2.3 in 2000, despite contraception being officially discouraged for some of that time. This was because women moved to cities and opted to have fewer babies.
25 The papers argue that the small, happy families portrayed on television contributed to this trend. Controlling for other factors, the arrival of Globo was associated with a decline of 0.6 percentage points in the probability of a woman giving birth in a given year. That is equivalent
30 to the drop in the birth rate associated with a woman having two extra years of schooling.

The effect on divorce was smaller, but noticeable. The researchers found that between 1975, when divorce was first discussed, and 1984, about one in five of the main characters in Globo soaps were divorced or separated, a higher percentage than in the real Brazil. These break-ups were not just a result of machismo: from the mid-1960s to the mid-1980s about 30% of female lead characters in *novelas* were unfaithful to their partners. The researchers found that the arrival of Globo in an area was associated with a rise of 0.1-0.2 percentage points in the share of women aged 15-49 who were divorced or separated. The authors reckon that watching “empowered” women having fun in Rio made other women (a few of them anyway) more independent.

Other research shows that divorce and lower fertility are linked to less domestic violence. So the influence of soaps may be far more positive than critics of their vapid claim. If Globo could now come up with a seductive *novela* about tax reform, its transformation of Brazil would be complete.

Source: *The Economist*, March 12th, 2009

1.1b Starry-eyed

For an ocean-loving nation, Japan has an odd way of showing its affection for the sea. Out of a fear of typhoons and tsunamis, and with an element of man-against-nature hubris, it has sealed much of its coastline in a girdle of concrete. Where children once scrambled over rocks to explore haunted caves, now they climb on cement tetrapods.

Hajime Kayanne, a science professor at the University of Tokyo, has a less intrusive approach to coping with the dangers from the sea. He calls it “eco-engineering”. It may represent the sort of “soft power” that Japan hopes to wield as its economic might is eclipsed by China.

Mr Kayanne’s special ingredient is not cement but star sand, the shells of a tiny single-celled organism, called *foraminifera*, found near coral reefs in Japan’s south – and taken in jars by tourists as a lucky charm. *Foraminifera* grow in the ocean, and when they die mix with coral to become the sediment that forms atolls. Mr Kayanne has found how to grow *foraminifera* artificially. He hopes this can help low-lying Pacific atolls survive if global warming raises sea levels.

65 Japan has plenty of self-interest in the endeavour. Mr Kayanne is
mainly focused on using star sand to raise the height of a tiny islet –
barely visible at low tide – called Okinotorishima, which is 1,740km
south of Tokyo.

Japan calls it an island and claims it as territory, which it says gives
it the right to an exclusive economic zone (EEZ) of 400,000 square
70 km – larger than Japan’s land mass itself. The surrounding sea is rich
in tuna, but is also militarily important to China. If rising sea levels
swamp the land (which is already encased in concrete), Japan loses
its claim to the EEZ.

So far, star sand has only been grown in the laboratory. It may not
75 even make it to Okinotorishima if lawyers decide that land propped
up by artificial star sand does not count as an island – in which case
it would not merit an EEZ.

But Mr Kayanne is also leading a Japanese project to create a
beach made out of home-grown star sand in Tuvalu, a South Pacific
80 atoll perilously close to sea level. He says up to three-quarters of the
island’s “body mass” – or sediment – is star sand, and he believes the
beach project will help regenerate the island naturally.

“Japan should be a world pioneer of this new eco-technology
against global warming,” he says, echoing a line from the election
85 manifesto of the new Japanese government. He admits the people
of Tuvalu have different ideas, however. “They say: ‘You come from
Japan. Why don’t you just build us a sea wall?’ ”

Source: *The Economist*, Oct. 15th, 2009

1.1c The Search for Sanctuary

A hundred or so orangutans have just returned from a day at “forest
school,” where they learn to find food, use tools and fear snakes.
90 Assembled on a lawn, they act much like any group of kids at recess
– congregating in small groups, clocking each other over the head,
holding hands, turning somersaults and adroitly climbing a jungle
gym. These apes are orphans, having lost their forest homes to
palm-oil plantations or their mothers to poachers who sell the babies
95 to an illegal pet trade. A few still bear the marks etched into their

necks from the months or years they spent in chains. The staff at the Nyaru Menteng orangutan sanctuary, in Central Kalimantan on the Indonesian side of Borneo, is painstakingly raising the apes and acting as surrogates for orangutan mothers who rear their young for up to eight years. Once they're old enough and wise enough to survive on their own, the apes move to islands that serve as halfway houses until the animals can be placed back into the jungle.

The orphaned orangutans, however, may have nowhere to go back to. As Borneo's rainforests continue to be felled at an alarming rate, suitable habitat is increasingly hard to come by. Roughly half of Borneo's forests have already been demolished over the past few decades, the trees turned to wood products and much of the land transformed into plantations for palm-oil trees. Orphanages are now scrambling for creative ways of financing the acquisition of land, while there's still some left.

The land squeeze is a direct consequence of the growing palm-oil trade. Palm oil is the world's most popular edible oil, used in products ranging from Cadbury's chocolate to Newman's Own cookies to Dove cleansers to biodiesel. Indonesia and Malaysia, which share the island of Borneo, together provide nearly 90 percent of the global supply. More than eight million acres of forest in Kalimantan are currently being razed or are allocated for palm plantations, according to a recent study. Resident animals are left homeless, particularly if the surrounding forest has already been destroyed. Workers hired to clear the land often kill the mother apes and sell the babies on the black market as pets. A small fraction are rescued by groups like the Borneo Orangutan Survival Foundation, which operates Nyaru Menteng.

The forest destruction has swelled the orphanage's ranks to more than 600 orangutans, even though the facility was initially designed to hold about 100. Although about 50,000 orangutans live on Borneo, most of them occupy unprotected areas, where they're vulnerable to development. The number of animals is declining very fast.

Conservationists have been trying to hold the line by defending existing strongholds. For instance, several groups, including WWF and the Frankfurt Zoological Society, recently banded together to protest a logging operation planned by two companies, Sinar Mas and Asia Pulp and Paper, which would destroy Sumatran rainforest where 100 rehabilitated orangutans have been released since 2002.

135 Wildlife organizations are also looking for new land, but there's not
much that's suitable. Orangutan habitat must be big enough to sustain
enough animals to breed and thrive. "If you find good lowland forest,
which we can't because it's almost all gone here, you could put about
two orangutans per square kilometer," said Lone Dröscher Nielsen,
140 the Danish director of Nyaru Menteng who has run the sanctuary for
more than 12 years. "Where we're looking we can probably only put
one per square kilometer. But you have to have at least one per square
kilometer or they'll never meet each other."

Further complicating matters, scientists are reluctant to mix
wild orangutans with those that have been raised by humans.
145 The rehabilitated animals could transmit human diseases – like
tuberculosis or hepatitis – to wild communities, and could themselves
be vulnerable to parasites or infections carried by wild orangutans.
And then there's the risk of transmitting social conventions – actions
or forms of expression the animals picked up from people that might
150 somehow pollute the wild culture of orangutans, threatening their
survival. "Recent research has shown that much of orangutan behavior
is learned, and there even appear to be regional variations," said John
Burton, founder of the World Land Trust, at a forum at the Linnean
Society in London in April. That means conservation groups must
155 find forest that doesn't currently house any orangutans, which rules
out many existing protected areas.

These restrictions have forced conservation groups to consider
forests they might have overlooked in the past. Scientists have found,
for instance, that orangutans do well in secondary forests, in which
160 ancient trees have been logged out and newer trees have grown to
take their place. The apes can also survive in forests that are logged
according to sustainable management principles. Those orangutans
that have thrived in one sustainably logged Sabah forest for a dozen
years often move through palm-oil plantations to get from one patch of
165 forest to another. Therefore plantation owners should be persuaded to
create corridors of rainforest trees linking otherwise isolated forests.
Such an arrangement could greatly expand the available forest to the
apes. Even the ecosystems that are protected are interconnected. So
efforts cannot be focused on protected forests alone.

170 The biggest hurdle, though, is raising the money to acquire
land. Conservation groups are banding together with bankers and

entrepreneurs to explore market-based solutions. An international program called Reduced Emissions from Deforestation and Degradation, or REDD, could enable governments and local communities to profit from healthy forests by capitalizing on the forests' ability to store carbon dioxide, a greenhouse gas. A study found that conserving Borneo's forests through the selling of carbon credits could be more profitable than turning them into palm-oil farms.

A growing number of firms are using REDD strategies to protect Borneo's forests, with deals in the works on parcels of up to several hundred thousand acres. "Forest is needed, first and foremost," says Richard Zimmerman, director of Orangutan Outreach, a New York conservation organization. "But is there hope? Yes. There has to be hope."

Source: *Newsweek*, Aug. 17th, 2009

1.2 Suomenkieliset koulut:

Lue seuraava teksti ja vastaa sen pohjalta **lyhyesti suomeksi** kysymyksiin a–e. Kirjoita vastauksesi **selvällä käsialalla** kielikokeen vastauslomakkeen **A-puolelle**.

Svenska skolor:

Läs texten och ge sedan **ett kort svar på svenska** på frågorna a–e. Skriv svaren **med tydlig handstil på sida A** av svarsblanketten för språkproven.

Olive investigates ORGANIC

Confused about organics? Who can blame you. For decades the merits, or otherwise, of organic food production have divided opinion.

The latest spat in the war of words around organics came when a review of scientific research into the possible nutritional benefits of organic food, carried out for the Food Standards Agency (FSA), found "no important differences in the nutrition content, or any additional health benefits, of organic food when compared with conventionally produced food". The validity of this conclusion was hotly contested by the pro-organic lobby. It asserted that the FSA review did not include the results of a major European Union-funded study, which showed

that levels of a range of nutritionally desirable compounds (such as vitamins and antioxidants) were higher in organic crops.

- a) Mitä FSA:n tutkimus totesi, ja mitä vastustajat väittivät?
Vad konstaterade man i FSA-undersökningen, och vad påstod motståndarna?

Another criticism of the FSA report was that it chose to ignore the possible health benefits of avoiding pesticides. Some 300 pesticides are approved for use in conventional food production, including known carcinogens and nerve poisons. Government monitoring shows that residues of these potent chemicals turn up in roughly a third of all the food we eat and almost half of all our fruit and vegetables. Regulators claim such residues pose no risk to human health because we consume only tiny traces. Critics of the Government's pesticide safety testing question this judgment, arguing that it doesn't give enough weight to the possible cocktail effect of accumulated residues of chemicals in the modern diet.

- b) Mistä FSA:n raporttia myös arvostellaan, ja mitä päättäjät väittävät?
Vad kritiseras FSA-rapporten också för, och vad påstår beslutsfattarna?

Defenders of conventional agriculture, however, highlight the fact that pesticide use has led to an explosion in crop yields over the last 50 years, effectively halving the cost of food and eradicating hunger in some countries. This contrasts with organic food, which is often more expensive. They warn that if we relied on organic farmers, there wouldn't be enough to feed the world's booming population. The Department for the Environment, Food and Rural Affairs has found that organic crop output is "an average of 42% lower than on conventional farms".

- c) Mitä etuja perinteisestä maanviljelystä sanotaan olevan luomuviljelyyn verrattuna?
Vilka fördelar påstås det konventionella jordbruket ha, jämfört med det ekologiska?

(continued on page 12)

1.1a Soaps, sex and sociology

1. What is said about the *telenovelas*?
 - A They superficially tell about Brazilian society
 - B They portray average Brazilian families
 - C They present a glamorised picture of Brazil
2. Why did the military regime support television?
 - A It was an efficient way to spread propaganda
 - B It could provide entertainment for the people
 - C It could reach those people who couldn't read
3. What have the soap scriptwriters and directors tried to do?
 - A Include serious issues in the plots
 - B Increase TV ratings
 - C Give examples of social changes
4. What did the Inter-American Development Bank study?
 - A If Globo's *novelas* affected the views on births and divorces
 - B If Globo's expansion had any social consequences
 - C If Globo's soaps had spread evenly throughout the country
5. Why did the birth rate drop so dramatically between 1960 and 2000?
 - A The use of contraception was encouraged by the state
 - B There was migration from rural to urban environments
 - C Women preferred to live alone and have fewer babies
6. What was the role of Globo in the declining fertility trend?
 - A It supported the idea of more education
 - B Its *novelas* told about small contented families
 - C It broadcast programmes on family planning and giving birth
7. What is the connection between Globo *novelas* and divorce?
 - A It is thought to be quite considerable
 - B There are fewer divorces because of the *novelas*
 - C *Novelas* may encourage more women to divorce

1.1b Starry-eyed

8. Why is there a cement wall around much of Japan?
 - A To show the Japanese love of the sea
 - B To protect against natural dangers
 - C To make the coastline safer for children
9. What is special about Mr Kayanne's approach?
 - A He recruits tourists as helpers
 - B He uses organic material
 - C He relies on "soft power"
10. What is said about coral atolls?
 - A They are partly made of *foraminifera*
 - B Global warming poses an immediate threat
 - C Tourists are causing damage to them
11. Why is Japan interested in Okinotorishima?
 - A It might help in protecting Japan against China
 - B It would serve several Japanese interests
 - C It could prevent the sea levels from rising
12. Why may Japan's plans fail?
 - A The use of star sand could be found illegal
 - B Star sand is possibly too artificial
 - C The EEZ might forbid the use of star sand
13. What other project does Mr Kayanne have?
 - A To make artificial beaches
 - B To help star sand grow
 - C To build up the Tuvalu atoll

1.1c The Search for Sanctuary

14. What is said about young orangutans?
 - A They prefer to live in groups
 - B They act much like human children
 - C They like to pick fights
15. Why are the young orangutans at the sanctuary?
 - A They have lost both their parents
 - B They are victims of human greed
 - C Their homes have been destroyed
16. What do we learn about orangutan mothers?
 - A They fight fiercely for their young
 - B They take care of their young for quite a long time
 - C They never abandon their young
17. What other problem do young orangutans face?
 - A They hardly ever learn how to live in the jungle again
 - B There are very few proper rainforests left
 - C It's difficult for them to leave the sanctuaries
18. Why is palm oil so valuable to people?
 - A It has so many uses
 - B It can be used as gas in cars
 - C There is a limited supply of it
19. What is one of the problems the founding of palm plantations has caused?
 - A Orangutans' sanctuaries have become too crowded
 - B It has badly affected the balance of nature
 - C Native people are forced to give up their land
20. Why has the number of orangutans decreased so fast?
 - A There are too many natural enemies in the jungle
 - B There is too little food for them to survive
 - C There is too much development taking place

21. How have the conservationist organizations reacted?
 - A They have tried to fight against the logging companies
 - B They have insisted the companies withdraw from Sumatra
 - C They have threatened to make their protest widely known
22. What would be an advantage of good lowland forest?
 - A It allows more orangutans to live there than on less suitable land
 - B It gives suitable shelter for orangutans' breeding
 - C It offers enough food and protection to orangutans
23. What is one of the dangers in rehabilitating orangutans?
 - A It may be a health risk to wild animals
 - B Wild animals' habitat may be polluted
 - C They may attack wild animals
24. What allows orangutans to thrive in logged forest areas?
 - A There are enough plantations there
 - B The forests are not totally logged
 - C Enough food is provided for them
25. How could Borneo's government profit from selling carbon credits?
 - A It could get rid of palm plantations
 - B It could bring more employment into the country
 - C It could utilize the healthy forests

A common reason for choosing organic food is animal welfare. The growth of factory farming has led to the intensive farming of livestock, especially of pigs, poultry and dairy cows. Such animals live entirely indoors, unable to express natural instincts such as foraging for food, or suckling their young. Intensive animal production systems are not permitted under organic rules. Organically reared animals must always be free range and they must be kept in smaller herds or flocks with more space to encourage good health and minimize stress.

- d) Mikä muu ero perinteisen ja luonnonmukaisen maatalouden välillä vielä on?
Vilken skillnad finns det ytterligare mellan konventionell och ekologisk lantushållning?

Another major difference is that organic farming is based on encouraging natural soil fertility to produce healthy crops and animals. Pests are kept under control by encouraging beneficial natural predators and a mixture of different crops is grown in rotation to discourage the build-up of disease that can occur when only one crop is grown continuously in the same spot. A scientific review has found that there are more birds, butterflies, beetles, bats and wild flowers on organic farms than on non-organic farms.

Source: *olive*, December 2009

- e) Millä tavoin luomuviljelyssä vältetään tuholaiset?
På vilka sätt undviker man skadeinsekter i det ekologiska jordbruket?

2 GRAMMAR AND VOCABULARY

- 2.1** *Read the text carefully and for each item choose the alternative that best fits the context. Mark your answers (26–50) on the optical answer sheet in pencil.*

The Drovers

From the departure of the Romans to the making of the toll roads in the 18th century no main public highways were built in Britain. People paved some paths as local need arose or travelled the old trade routes. These paths 26 by lines of heavily burdened packhorses carrying essential goods.

But for centuries, the most important long-distance travellers were the drovers: Celts from Scotland and Wales bringing their highland cattle across England to be fattened in the grazing lands of the south-east before 27 at the London or Kent markets. They formed great cavalcades that blocked the way for other travellers for hours 28. Some parts of the droveways were also used to transport pigs, sheep, geese and turkeys, and these animals also had to travel great distances. But 29 the cattle drovers who were the aristocrats of their trade.

26. A use
B used
C used to
D were used

27. A sold
B being sold
C having sold
D to be sold

28. A on time
B for a time
C at a time
D with the times

29. A it was
B there was
C there were
D they were

Although there are wild tales of the drovers, 30 the exploits of the famous Highlander, Rob Roy, for the most part they were solid worthies, 31 their neighbours to carry money and important documents as well as cattle to the south, and bring home goods and news of the outside world to the remote farmsteads. They were the reporters of the time, 32 were the first to hear and tell of the outcome of the battle of Waterloo, and to whom the farmers' wives turned for information about the 33 fashions.

Above all they were responsible 34 the large sums of money which the cattle represented – on the hoof on the outward journey and homeward in hard cash. This wealth had to 35 through remote mountain passes haunted by highwaymen. The obvious solution was to devise a safe means of transferring cash without having to carry 36 about. So the history of banking is closely linked to the droving trade.

The more prosperous Welsh drovers had overnight accommodation at inns, the Highlanders lived more sparsely, sleeping out of doors with the cattle. For both, the

30. A as
B as if
C such as
D example
31. A trusting
B trusting in
C trusted
D trusted by
32. A that
B which
C who
D those
33. A late
B later
C latter
D latest
34. A for
B of
C to
D with
35. A take
B taking
C be taken
D have taken
36. A it
B that
C those
D –

first consideration was always the grazing of the beasts. 37 in most cases all traces of the old inns have completely disappeared, you can still track down 38 whereabouts by patches of bright grass, owing its lushness to centuries of manuring by the visiting herds. Halfpenny pastures they are often called, reflecting the cost of grazing per night per beast.

The cattle 39 often polled, partly because it is easier to handle beasts without horns, and partly because of an old folk belief that the growth of horns took up 40 of the vital juices that should go to the building up of flesh. The cattle had to be shod for their long journey. As cattle have cloven hoofs, the shoe for each hoof had to be made in two parts. It takes a special skill and enormous strength to shoe an ox. But one shoeing would not last 41 long trip. Local blacksmiths did well to have their forges sited on the drove routes. Sometimes, too, a smith with a portable forge 42 accompany the droves on part of the route, to replace cast shoes or to deal with any animal which showed signs of becoming lame.

43 all other trades and crafts, the drovers developed their own folklore and

37. A Although
B Because
C Whereas
D When
38. A its
B their
C the
D some
39. A was
B were
C has been
D have been
40. A any
B each
C none
D some
41. A all
B all of
C the whole of
D the whole of the
42. A need to
B were to
C should
D would
43. A Among
B As
C Like
D With

superstition and 44 their own work songs and ballads as they rode or tramped with the cattle across the hills. Rowan was the plant that above all others brought good luck, warding off supernatural evil and natural accidents. Until the 17th century packs of wolves still roamed the wilder moorlands, but probably a 45 most anxious moments came when water 46 crossed. Often there were no boats, and beasts, men and horses had to swim the straits or make a hurried journey across the sands at low water.

The regular 47 of the journey – an average of two miles an hour – must also have been a consideration for 48 many people who travelled with the drovers 49 have companionship and some measure of safety on their journeys. The sons of rich landowners went to London with the drovers to see the world; other lads from more modest backgrounds were sent by their fathers to take up apprenticeship in the English cities. And they 50 were used by the drovers to ward off highwaymen and bandits.

Source: Shirley Toulson, *The Drovers*, 1980

44. A sing
B singing
C sang
D sung
45. A drover
B drovers
C drover's
D drovers'
46. A has to be
B had to be
C had been
D was to
47. A fastness
B pace
C rapidity
D rates
48. A a
B an
C the
D –
49. A for
B in order
C in order to
D so that
50. A in favour of
B in return
C on the other hand
D the other way round

2.2 Translate clauses (1)–(6) *into natural English*. Write your answers in the given order **on side B** of the answer sheet. Write each translation on a separate line. Please write *clearly*.

Dear Aunt Annie,

I'm sorry for not having written to you earlier, but I've been busy with, you know, my studies and, well,

- (1) *ainahan niitä syitä löytyy.* (2 p.)
ursäkter finns det alltid.

But anyway, I wish to thank you SO MUCH for the nice birthday present. I knew I could rely on my favourite aunt to know

- (2) *mikä olisi sisarenpojalle mieluista.* (2 p.)
vad en systerson skulle tycka om.

The money was just what I needed to make my dream come true.

- (3) *Merimatka Seawindillä oli upea kokemus,* (2 p.)
Sjöresan på Seawind var en fantastisk upplevelse,

and I wasn't seasick at all. The hotel in Scotland was just right and the food was OK, too.

- (4) *Lainasin pyörän ja ajoin kymmeniä maileja yhden* (3 p.)
kaverin kanssa.
Jag lånade en cykel och åkte tiotals mil tillsammans
med en kompis.

We found some nice shells for your collection and almost drowned because we forgot to keep an eye on high tide. (Please don't tell Mom!)

- (5) *Emme kuitenkaan menneet uimaan, koska siellä oli* (3 p.)
niin kylmä.
Vi gick ändå inte och badade, för det var så kallt där.

I saw sheep everywhere and some men in kilts but never heard any bagpipes. Well, I'm home again and can only dream of those days with my nose in my books here at college.

- (6) *Äidiltä terveisiä, ja nähdään taas jouluna.
Hälsningar från mamma, och vi ses igen i jul.* (3 p.)

With fondest love,
David

total (15 p.)

3 PRODUCTION

*Write a composition of between 150 and 250 words on one of the following topics. Please write **clearly** on the notebook paper (konseptipaperi/konceptpapper) provided. Follow the guidance. Count the number of words in your composition and write it at the end.*

1. Eat in, eat out, eat away

When you want to celebrate something with your friends, do you yourself prepare the food or do you prefer some other arrangement? Why? What does your “menu” consist of?

2. Money down the drain

Consuming is the trend of the day. What are real necessities? What could we do without? How could we save money?

3. Shopping on Sundays

Is it necessary to have shops open all days of the week? What advantages or disadvantages has it brought? Write a letter to the editor of a newspaper.

4. Programming children 24/7

Today, parents seem to think that children have to be constantly programmed with various activities and hobbies. What is your idea of “quality time” for children?

KOKEEN PISTEITYS / POÄNGSÄTTNINGEN AV PROVET

Tehtävä	Osioiden määrä	Pisteitys	Paino-kerroin*	Enint.	Arvostelulomakkeen sarake
Uppgift	Antal deluppgifter	Poängsättning	Koefficient*	Max.	Kolumn på bedömningsblanketten
1.1a–c	25 x	1/0 p.	x 2	50 p.	1
1.2	5 x	2–0 p.	x 2	20 p.	2
2.1	25 x	1/0 p.	x 1	25 p.	3
2.2	6 x	2/3–0 p.	x 1	15 p.	4
3				99 p.	7
Yht./Tot.				209 p.	

* Painotus tapahtuu lautakunnassa.
Viktningen görs av nämnden.

Appendix 4



Eksamen

22.5.2009

SPR3012 Engelskspråkleg litteratur og kultur /
Engelskspråklig litteratur og kultur
Elevar/Elever, Privatistar/Privatister

Nynorsk/Bokmål

Nynorsk

Eksamensinformasjon	
Eksamenstid:	5 timar
Hjelpemiddel:	Alle hjelpemiddel er tillatne. Unntak er Internett og andre verktøy som tillet kommunikasjon. For norsk, samisk, finsk som 2. språk og framandspråka er heller ikkje omsetjingsprogram tillatne.
Bruk av kjelder:	Alle kjelder som blir nytta til eksamen, skal først opp på ein slik måte at lesaren kan finne fram til kjelda. Du må føre opp forfattar og heile tittelen både på lærebøker og annan litteratur. Dersom du har med deg utskrift eller sitat frå nettsider, skal heile adressa og dato for nedlastinga først opp. Det er til dømes ikkje tilstrekkeleg med www.wikipedia.no
Vedlegg:	1 <i>Never Let Me Go</i> 2 "What is mobbing?"
Informasjon om vurderinga:	<p>På eksamensdagen må du skrive ein sjølvstendig tekst som er eit relevant svar på den valde oppgåva. Godt språk, logisk struktur og sjølvstendig bruk av kjelder blir også vektlagde i vurderinga.</p> <p>Oppgåvesettet har tre hovuddelar: oppgåve 1, oppgåve 2 og oppgåve 3. Du skal svare på alle tre delane. I oppgåve 3 skal du svare på éi oppgåve, 3a, 3b, 3c eller 3d.</p> <p>Eit godt svar på oppgåve 1 viser at du kan drøfte språklege særpreg ved tekstar i ulike sjangrar. Bruk av presis terminologi og relevante eksempel frå tekstane vil bli honorerert.</p> <p>Eit godt svar på oppgåve 2 viser at du kan tolke ein litterær tekst i samfunnsperspektiv, og at du kan drøfte aktuelle spørsmål i det internasjonale nyheitsbiletet.</p> <p>Alle deloppgåvene i oppgåve 3 måler kompetansen til å skrive tekstar med god struktur og samanheng om allmenne, faglege eller litterære emne.</p> <p>Eit godt svar på oppgåve 3a viser at du kan analysere og drøfte litteratur.</p> <p>Eit godt svar på oppgåve 3b viser at du kan produsere ein tekst i sjangeren <i>novelle</i> med klart innhald og formålstenleg stil, struktur og språkbruk.</p> <p>Eit godt svar på oppgåve 3c viser at du kan tolke ein tekst frå ein bestemt litteraturhistorisk epoke.</p>

	Eit godt svar på oppgåve 3d viser at du kan tolke litterære tekstar, eventuelt analysere og vurdere ein film.
Kjelder:	<ol style="list-style-type: none"> 1 "Vedlegg 1": Ishiguro, K. 2005. <i>Never Let Me Go</i>. faber and faber: London. 2 "Vedlegg 2": "What is mobbing?" http://www.bullyonline.org/workbully/mobbing.html (08.10.2008). 3 Giovanni, N. 1968. "The Funeral of Martin Luther King Jr.". Black Feeling, Black Talk, Black Judgement. New York: W. Morrow & Co. 4 Wordsworth, W. "The Solitary Reaper" http://www.bartleby.com/41/388.html (07.10.2008).

Bokmål

Eksamensinformasjon	
Eksamenstid:	5 timer
Hjelpemidler:	Alle hjelpemidler er tillatt. Unntak er Internett og andre verktøy som tillater kommunikasjon. For norsk, samisk, finsk som 2. språk og fremmedspråkene er heller ikke oversettelsesprogrammer tillatt
Bruk av kilder:	Alle kilder som blir brukt til eksamen skal oppgis på en slik måte at leseren kan finne fram til kilden. Du må oppgi forfatter og hele tittelen både på lærebøker og annen litteratur. Dersom du har med deg utskrift eller sitat fra nettsider, skal hele adressen og nedlastingsdato oppgis. Det er f. eks. ikke tilstrekkelig med www.wikipedia.no .
Vedlegg:	1 Never Let Me Go 2 What is mobbing?"
Informasjon om vurderingen:	<p>På eksamensdagen må du skrive en selvstendig tekst som er et relevant svar på den valgte oppgaven. Godt språk, logisk struktur og selvstendig brukt av kilder blir også vektlagt i vurderingen.</p> <p>Oppgavesettet har tre hoveddeler: oppgave 1, oppgave 2 og oppgave 3. Du skal svare på alle tre delene. I oppgave 3 skal du svare på én oppgave, 3a, 3b, 3c eller 3d.</p> <p>Et godt svar på oppgave 1 viser at du kan drøfte språklig særpreg ved tekster i ulike sjangrer. Bruk av presis terminologi og relevante eksempler fra tekstene vil bli honorert.</p> <p>Et godt svar på oppgave 2 viser at du kan tolke en litterær tekst i et samfunnsmessig perspektiv og at du kan drøfte aktuelle spørsmål i det internasjonale nyhetsbildet.</p> <p>Alle deloppgavene i oppgave 3 måler kompetansen til å skrive tekster med god struktur og sammenheng om allmenne, faglige eller litterære emner.</p> <p>Et godt svar på oppgave 3a viser at du kan analysere og drøfte litteratur.</p> <p>Et godt svar på oppgave 3b viser at du kan produsere en tekst i sjangeren <i>novelle</i> med klart innhold og hensiktsmessig stil, struktur og språkbruk.</p> <p>Et godt svar på oppgave 3c viser at du kan tolke en tekst fra en bestemt litteraturhistorisk epoke.</p> <p>Et godt svar på oppgave 3d viser at du kan tolke litterære tekster, eventuelt analysere og vurdere en film.</p>

Kilder:	<ol style="list-style-type: none"> 1 "Vedlegg 1": Ishiguro, K. 2005. <i>Never Let Me Go</i>. faber and faber: London. 2 "Vedlegg 2": "What is mobbing?" http://www.bullyonline.org/workbully/mobbing.htm (08.10.2008). 3 Giovanni, N. 1968. "The Funeral of Martin Luther King Jr.". <i>Black Feeling, Black Talk, Black Judgement</i>. New York: W. Morrow & Co. 4 Wordsworth, W. "The Solitary Reaper" http://www.bartleby.com/41/388.html (07.10.2008).
----------------	--

Oppgave 1 / Oppgave 1

Answer **both** a) and b).

a)

Write a text in which you answer the following question:

What is the genre of Never Let Me Go in "Vedlegg 1", and what is the genre of "What is mobbing?" in "Vedlegg 2"? Give reasons for your answer.

b)

The writers use language in different ways in the two extracts.

Discuss some of these differences in language use in the two extracts. Illustrate the differences with specific examples from both texts.

Oppgave 2 / Oppgave 2

Write two or three paragraphs in which you discuss how the poem "The Funeral of Martin Luther King, Jr." below expresses some of the feelings evoked by the election of Barack Obama as US President in 2008.

The Funeral of Martin Luther King, Jr.

His headstone said
FREE AT LAST, FREE AT LAST
But death is a slave's freedom
We seek the freedom of free men
And the construction of a world
Where Martin Luther King could have lived
And preached non-violence.

The poem was written by the female American poet Nikki Giovanni in 1968.

Oppgave 3 / Oppgave 3

Choose **one** of the alternatives below, a), b), c) **or** d):

a)

On the basis of the extract from the novel Never Let Me Go in "Vedlegg 1", write a character analysis of Tommy and Kathy. Comment on how the point of view is important for your understanding of these two characters.

b)

Write a short story in which you develop the relationship between Kathy and Tommy, the two main characters in the extract from the novel Never Let Me Go in "Vedlegg 1".

c)

Write an analysis of the poem below pointing out its typical Romantic features.

The Solitary Reaper

BEHOLD her, single in the field,
Yon solitary Highland Lass!
Reaping and singing by herself;
Stop here, or gently pass!
Alone she cuts and binds the grain,
And sings a melancholy strain;
O listen! for the Vale profound
Is overflowing with the sound.

No Nightingale did ever chaunt
More welcome notes to weary bands
Of travellers in some shady haunt,
Among Arabian sands:
A voice so thrilling ne'er was heard
In spring-time from the Cuckoo-bird,
Breaking the silence of the seas
Among the farthest Hebrides.

Will no one tell me what she sings?--
Perhaps the plaintive numbers flow
For old, unhappy, far-off things,
And battles long ago:
Or is it some more humble lay,
Familiar matter of to-day?
Some natural sorrow, loss, or pain,
That has been, and may be again?

Whate'er the theme, the Maiden sang
As if her song could have no ending;
saw her singing at her work,
And o'er the sickle bending:--
I listened, motionless and still;
And, as I mounted up the hill
The music in my heart I bore,
Long after it was heard no more.

William Wordsworth (1770–1850)
<http://www.bartleby.com/41/388.html> (07 Nov. 2008)

d)

Write an essay in which you compare the way cruelty in human relationships is explored in the extract from Never Let Me Go in "Vedlegg 1" and in another literary work or a film you have studied.

The text below is an extract from the novel *Never Let Me Go*. Kathy, Ruth and Tommy are all pupils at the English boarding school Hailsham. Kathy looks back on her school days, remembering Tommy and a special event that took place on the sports field.

Never Let Me Go

Or maybe I'm remembering it wrong. Maybe even then, when I saw Tommy rushing about that field, undisguised delight on his face to be accepted back in the fold again, about to play the game at which he so excelled, maybe I did feel a little stab of pain. What I do remember is that I noticed Tommy was wearing the light blue polo shirt he'd got in the Sales the previous month — the one he was so proud of. I remember thinking: "He's really stupid, playing football in that. It'll get ruined, then how's he going to feel?" Out loud, I said, to no one in particular: "Tommy's got his shirt on. His favourite polo shirt."

I don't think anyone heard me, because they were all laughing at Laura — the big clown in our group —mimicking one after the other the expressions that appeared on Tommy's face as he ran, waved, called, tackled. The other boys were all moving around the field in that deliberately languorous way they have when they're warming up, but Tommy, in his excitement, seemed already to be going full pelt. I said, louder this time: "He's going to be so sick if he ruins that shirt." This time Ruth heard me, but she must have thought I'd meant it as some kind of joke, because she laughed half-heartedly, and then made some quip of her own.

Then the boys had stopped kicking the ball about and were standing in a pack in the mud, their chests gently rising and falling as they waited for the team picking to start. The two captains who emerged were from Senior 3, though everyone knew Tommy was a better player than any of that year. They tossed for first pick, and then the one who'd won stared at the group.

"Look at him," someone behind me said. "He's completely convinced he's going to be first pick. Just look at him!"

There was something comical about Tommy at that moment, something that made you think, well, yes, if he's going to be that daft, he deserves what's coming. The other boys were all pretending to ignore the picking process, pretending they didn't care where they came in the order. Some were talking quietly to each other, some re-tying their laces, others just staring down at their feet as they trammelled the mud. But Tommy was looking eagerly at the Senior 3 boy, as though his name had already been called.

Laura kept up her performance all through the team picking, doing all the different expressions that went across Tommy's face: the bright eager one at the start; the puzzled concern when four picks had gone by and he still hadn't been chosen; the hurt and panic as it began to dawn on him what was really going on. I didn't keep glancing round at Laura, though, because I was watching Tommy; I only knew what she was doing because the others kept laughing and egging her on. Then when Tommy was left standing alone, and the boys all began sniggering, I heard Ruth say:

"It's coming. Hold it. Seven seconds. Seven, six, five . . ."

She never got there. Tommy burst into thunderous bellowing, and the boys, now laughing openly, started to run off towards the South Playing Field. Tommy took a few strides after them — it was hard to say whether

his instinct was to give angry chase or if he was panicked at being left behind. In any case he soon stopped and stood there, glaring after them, his face scarlet. Then he began to scream and shout, a nonsensical jumble of swear words and insults.

We'd all seen plenty of Tommy's tantrums by then, so we came down off our stools and spread ourselves around the room. We tried to start up a conversation about something else, but there was Tommy going on and on in the background, and although at first we just rolled our eyes and tried to ignore it, in the end — probably a full ten minutes after we'd first moved away — we were back up at the windows again.

The other boys were now completely out of view, and Tommy was no longer trying to direct his comments in any particular direction. He was just raving, flinging his limbs about, at the sky, at the wind, at the nearest fence post. Laura said he was maybe "rehearsing his Shakespeare." Someone else pointed out how each time he screamed something he'd raise one foot off the ground, pointing it outwards, "like a dog doing a pee." Actually, I'd noticed the same foot movement myself, but what had struck me was that each time he stamped the foot back down again, flecks of mud flew up around his shins. I thought again about his precious shirt, but he was too far away for me to see if he'd got much mud on it.

"I suppose it is a bit cruel," Ruth said, "way they always work him up like that. But it's his own fault. If he learnt to keep his cool, they'd leave him alone."

"They'd still keep on at him," Hannah said. "Graham K.'s temper's just as bad, but that only makes them all the more careful with him. The reason they go for Tommy is because he's a layabout."

Then everyone was talking at once, about how Tommy never even tried to be creative, about how he hadn't even put anything in for the Spring Exchange. I suppose the truth was, by that stage, each of us was secretly wishing a guardian would come from the house and take him away. And although we hadn't had any part in this latest plan to rile Tommy, we had taken out ringside seats, and we were starting to feel guilty. But there was no sign of a guardian, so we just kept swapping reasons why Tommy deserved everything he got. Then when Ruth looked at her watch and said even though we still had time, we should get back to the main house, nobody argued.

Tommy was still going strong as we came out of the pavilion. The house was over to our left, and since Tommy was standing in the field straight ahead of us, there was no need to go anywhere near him. In any case, he was facing the other way and didn't seem to register us at all. All the same, as my friends set off along the edge of the field, I started to drift over towards him. I knew this would puzzle the others, but I kept going — even when I heard Ruth's urgent whisper to me to come back.

I suppose Tommy wasn't used to being disturbed during his rages, because his first response when I came up to him was to stare at me for a second, then carry on as before. It was like he was doing Shakespeare and I'd come up onto the stage in the middle of his performance. Even when I said: "Tommy, your nice shirt. You'll get it all messed up," there was no sign of him having heard me.

So I reached forward and put a hand on his arm. Afterwards, the others thought he'd meant to do it, but I was pretty sure it was unintentional. His arms were still flailing about, and he wasn't to know I was about to put out my hand. Anyway, as he threw up his arm, he knocked my hand aside and hit the side of my face. It didn't hurt at all, but I let out a gasp, and so did most of the girls behind me.

That's when at last Tommy seemed to become aware of me, of the others, of himself, of the fact that he was there in that field, behaving the way he had been, and stared at me a bit stupidly.

"Tommy," I said, quite sternly. "There's mud all over your shirt."

"So what?" he mumbled. But even as he said this, he looked down and noticed the brown specks, and only just stopped himself crying out in alarm. Then I saw the surprise register on his face that I should know about his feelings for the polo shirt.

"It's nothing to worry about," I said, before the silence got humiliating for him. "It'll come off. If you can't get it off yourself, just take it to Miss Jody."

He went on examining his shirt, and then said grumpily: "It's nothing to do with you anyway."

He seemed to regret immediately this last remark and looked at me sheepishly, as though expecting me to say something comforting back to him. But I'd had enough of him by now, particularly with the girls watching — and for all I knew, any number of others from the windows of the main house. So I turned away with a shrug and rejoined my friends.

Ruth put an arm around my shoulders as we walked away. "At least you got him to pipe down," she said. "Are you okay? Mad animal."

Kazuo Ishiguro

What is mobbing?

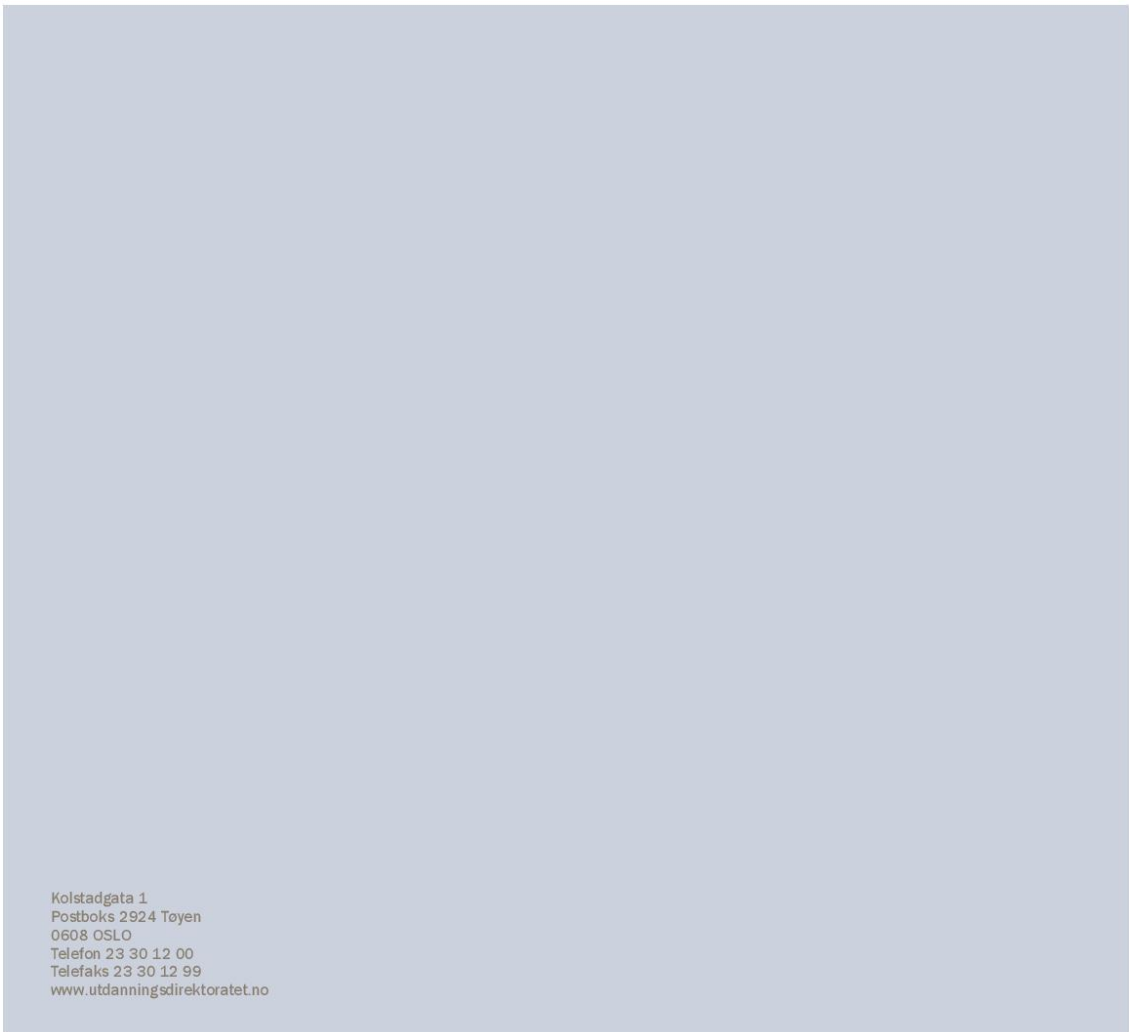
The word bullying is used to describe a repeated pattern of negative intrusive violational behaviour against one or more targets and comprises constant trivial nit-picking criticism, refusal to value and acknowledge, undermining, discrediting and a host of other behaviours which are defined on the page "What is bullying?"

The word *mobbing* is preferred to bullying in continental Europe and in those situations where a target is selected and bullied (mobbed) by a group rather than by one individual.

In a mobbing situation a ringleader incites supporters, cohorts, copycats and unenlightened, inexperienced, immature or emotionally needy individuals with poor values to engage in adversarial interaction with the selected target. The ringleader, or chief bully, gains gratification from encouraging others to engage in adversarial interaction with the target. Many people use the word "mobbing" to describe this pack attack by many on one individual.

One aspect of psychopathic bullies is that they home in on Wannabe types – non-psychopathic lesser bullies – and then empower these individuals to gain the positions of power and authority they crave. Once installed, the Wannabe's lack of competence makes them dependent on the chief psychopath, which means they become unwitting but willing compliant puppets.

Based on "What is mobbing?" at
<http://www.bullyonline.org/workbully/mobbing.htm>



Kolstadgata 1
Postboks 2924 Tøyen
0608 OSLO
Telefon 23 30 12 00
Telefaks 23 30 12 99
www.utdanningsdirektoratet.no

Appendix 5



Eksamen

27.05.2010

SPR3012

Engelskspråkleg litteratur og kultur/

Engelskspråklig litteratur og kultur

Elevar/Elever, Privatistar/Privatister

Nynorsk/Bokmål

Nynorsk

Eksamensinformasjon	
Eksamenstid:	5 timar
Hjelpemiddel:	Alle hjelpemiddel er tillatne. Unntak er Internett og andre verktøy som tillèt kommunikasjon. For norsk, samisk, finsk som 2. språk og framandspråka er heller ikkje omsetjingsprogram tillatne.
Bruk av kjelder:	Kjelder er tekstar som er tilgjengelege for andre, dvs. publiserte tekstar. Dersom det er aktuelt for deg å bruke kjelder i svaret ditt – anten fordi oppgåva krev det, eller fordi du vel å bruke kjelder, må desse førast opp på ein etterretteleg måte. Det finst ulike måtar å føre opp kjelder på. Det vesentlege er at alle kjelder som blir brukte til eksamen, skal førast opp på ein slik måte at lesaren kan finne fram til dei. Du skal føre opp forfattar og fullstendig tittel på så vel lærebøker som annan litteratur. Dersom du bruker utskrift eller sitat frå nettsider, skal du føre opp nøyaktig nettadresse og nedlastingsdato.
Vedlegg:	<ol style="list-style-type: none"> 1. "In Goya's Greatest Scenes We Seem to See ..." av Lawrence Ferlinghetti 2. Bilete 1: "The Shootings of May 3, 1808" av Francisco Goya Bilete 2: "Saturn Devouring One of His Sons" av Francisco Goya Bilete 3: "Traffic" 3. "Misunderstood" av Pete Townshend 4. Utdrag frå <i>The Scarlet Letter</i> av Nathaniel Hawthorne
Informasjon om oppgåva:	<p>Oppgåvesettet har tre hovuddelar: oppgåve 1, oppgåve 2 og oppgåve 3. Du skal svare på alle tre oppgåvene. I oppgåve 1 skal du svare på begge deloppgåvene a) og b). I oppgåve 3 skal du svare på éi deloppgåve: 3a), 3b), 3c) eller 3d).</p> <p>Kompetansemål som vert prøvde:</p> <p>Alle oppgåvene måler kompetanse i å produsere tekstar i ulike sjangrar med klart innhald, føremålstenleg stil, god struktur, presis språkbruk og rikt og nyansert ordforråd til å kommunisere om litteratur og kultur, samt kompetanse i å bruke føremålstenleg og situasjonstilpassa språk i skriftlege sjangrar.</p> <p>Oppgåve 1:</p> <ul style="list-style-type: none"> • drøfte samanhengen mellom form, innhald og stilnivå i ein tekst • drøfte språkleg krevjande framstillingar med fagleg innhald • bruke føremålstenleg språk tilpassa ein film-/litteraturkritikk • drøfte ein sakprosa tekst frå engelskspråkleg kultur <p>Oppgåve 2:</p> <ul style="list-style-type: none"> • drøfte samanhengen mellom form og innhald i ein annonse • analysere eit bilete • drøfte aktuelle spørsmål i det internasjonale nyheitsbiletet

	<p>Oppgave 3a: • tolke dikt og andre kulturuttrykk i eit kulturhistorisk og samfunnsmessig perspektiv</p> <p>Oppgave 3b: • produsere tekstar i ulike sjangrar</p> <p>Oppgave 3c og 3d:</p> <ul style="list-style-type: none"> • tolke og analysere skjønnlitterære tekstar • meistre terminologi til å analysere skjønnlitteratur
Kjelder:	<p>Oppgave 1: "Text": http://www.thepequod.org.uk/essays/litcrit/visualpo.htm</p> <p>Oppgave 2: http://adland.tv/ooh/amnesty-international-ads-show-whats-happening</p> <p>Oppgave 3a: Vedlegg 1: "In Goya's Greatest Scenes We Seem to See ..." frå Ferlinghetti, L. 1958. <i>A Coney Island of the Mind, Poems by Lawrence Ferlinghetti</i>. New Directions Paperbook: pp.9-10.</p> <p>Vedlegg 2: Bilete 1: "The Shootings of May 3, 1808" av Francisco Goya: history.hanover.edu/courses/art/goyamy3.html (27.11.2009)</p> <p>Vedlegg 2: Bilete 2: "Saturn Devouring One of His Sons" av Francisco Goya: http://www.online-literature.com/forums/showthread.php?t=23757&page=10 (27.11.2009)</p> <p>Vedlegg 2: Bilete 3: "Traffic": http://multivu.prnewswire.com/mnr/affinion/27862/images/27862-hi-traffic.jpg</p> <p>Oppgave 3b: Vedlegg 3: "Misunderstood" av Pete Townshend: http://www.lyricsdepot.com/pete-townshend/misunderstood/html</p> <p>Oppgave 3c: Vedlegg 4: Utdrag frå <i>The Scarlet Letter</i> av Nathaniel Hawthorne: http://www.online-literature.com/hawthorne/scarletletter/ (26.11.2009)</p> <p>Oppgave 3d: Utdanningsdirektoratet</p>
Informasjon om vurderinga:	<p>På eksamensdagen må du skrive ein sjølvstendig tekst som er eit relevant svar på oppgåva. Godt språk, logisk struktur og sjølvstendig bruk av kjelder blir også vektlagde i vurderinga.</p>

Bokmål

Eksamensinformasjon	
Eksamenstid:	5 timer
Hjelpemidler:	Alle hjelpemidler er tillatt. Unntak er Internett og andre verktøy som tillater kommunikasjon. For norsk, samisk, finsk som 2. språk og fremmedspråkene er heller ikke oversettelsesprogrammer tillatt.
Bruk av kilder:	Kilder er tekster som er tilgjengelige for andre, dvs. publiserte tekster. Dersom det er aktuelt for deg å bruke kilder i din besvarelse – enten fordi oppgaven krever det, eller fordi du velger å bruke kilder, må disse oppgis på en etterrettelig måte. Det finnes ulike måter å oppgi kilder på. Det vesentlige er at alle kilder som blir brukt til eksamen, skal oppgis på en slik måte at leseren kan finne fram til dem. Du skal oppgi forfatter og fullstendig tittel på så vel lærebøker som annen litteratur. Dersom du bruker utskrift eller sitat fra nettsider, skal du oppgi nøyaktig nettadresse og nedlastingsdato.
Vedlegg:	1. "In Goya's Greatest Scenes We Seem to See ..." av Lawrence Ferlinghetti 2. Bilde 1: "The Shootings of May 3, 1808" av Francisco Goya Bilde 2: "Saturn Devouring One of His Sons" av Francisco Goya Bilde 3: "Traffic" 3. "Misunderstood" av Pete Townshend 4. Utdrag fra <i>The Scarlet Letter</i> av Nathaniel Hawthorne
Informasjon om oppgaven:	Oppgavesettet har tre hoveddeler: oppgave 1, oppgave 2 og oppgave 3. Du skal svare på alle tre oppgavene. I oppgave 1 skal du svare på begge deloppgavene, a) og b). I oppgave 3 skal du svare på én deloppgave, 3a), 3b), 3c) eller 3d). Kompetansemål som prøves: Alle oppgavene måler kompetanse i å produsere tekster i ulike sjangere med klart innhold, hensiktsmessig stil, god struktur, presis språkbruk og et rikt og nyansert ordforråd til å kommunisere om litteratur og kultur, samt kompetanse i å bruke hensiktsmessig og situasjonstilpasset språk i skriftlige sjangere. Oppgave 1: <ul style="list-style-type: none"> • drøfte sammenhengen mellom form og innhold og stilnivå i en tekst • drøfte språklig krevende framstillinger med faglig innhold • bruke hensiktsmessig språk tilpasset en film-/litteraturkritikk • drøfte en sakprosaetekst fra engelskspråklig kultur Oppgave 2: <ul style="list-style-type: none"> • drøfte sammenhengen mellom form og innhold i en annonse • analysere et bilde

	<ul style="list-style-type: none"> • drøfte aktuelle spørsmål i det internasjonale nyhetsbildet <p>Oppgave 3a: • tolke dikt og andre kulturuttrykk i et kulturhistorisk og samfunnsmessig perspektiv</p> <p>Oppgave 3b: • produsere tekster i ulike sjangere</p> <p>Oppgave 3c og 3d:</p> <ul style="list-style-type: none"> • tolke og analysere skjønnlitterære tekster • beherske terminologi til å analysere skjønnlitteratur
Kilder:	<p>Oppgave 1: "Text": http://www.thepequod.org.uk/essays/litcrit/visualpo.htm</p> <p>Oppgave 2: http://adland.tv/oooh/amnesty-international-ads-show-whats-happening</p> <p>Oppgave 3a: Vedlegg 1: "In Goya's Greatest Scenes We Seem to See ..." fra Ferlinghetti, L. 1958. <i>A Coney Island of the Mind, Poems by Lawrence Ferlinghetti</i>. New Directions Paperbook: pp.9-10. Vedlegg 2: Bilde 1: "The Shootings of May 3, 1808" av Francisco Goya: http://history.hanover.edu/courses/art/goyamy3.html (27.11.2009) Vedlegg 2: Bilde 2: "Saturn Devouring One of His Sons" av Francisco Goya: http://www.online-literature.com/forums/showthread.php?t=23757&page=10 (27.11.2009) Vedlegg 2: Bilde 3: "Traffic": http://multivu.prnewswire.com/mnr/affinion/27862/images/27862-hi-traffic.jpg</p> <p>Oppgave 3b: Vedlegg 3: "Misunderstood" av Pete Townshend: http://www.lyricsdepot.com/pete-townshend/misunderstood/html</p> <p>Oppgave 3c: Vedlegg 4: Utdrag fra <i>The Scarlet Letter</i> av Nathaniel Hawthorne: http://www.online-literature.com/hawthorne/scarletletter/ (26.11.2009)</p> <p>Oppgave 3d: Utdanningsdirektoratet</p>
Informasjon om vurderingen:	<p>På eksamensdagen må du skrive en selvstendig tekst som er et relevant svar på oppgaven. Godt språk, logisk struktur og selvstendig bruk av kilder blir også vektlagt i vurderingen.</p>

Oppgave 1 / Oppgave 1

Read the text in the frame and answer questions a) **and** b) below.

Text:

1 The theatre is marked by its ability to encompass extremes of dialogic and visual show. The mime does not require speech in order to deliver effective drama but alternatively a theatre set may be almost totally clear, with the prime focus of the audience being on dialogue. Samuel Beckett, at the forefront of modernist theatre, offers examples at both ends of this spectrum; his *Act Without Words* is a remarkable mime, whilst *Play* promotes the dialogue over the setting, with lights and urns being used, passively, as means to highlight disembodied voices. Film has less extreme options since it is a primarily visual medium. The silent film is still a film, the screen-less film is not. Shakespeare's plays are positioned towards the dialogic mode although they retain a high potential for visual elements. Physical setting is never fundamental to the structure of the drama and the absence of stage directions creates an interpretative space for producers to fill. It is language, rather than external objects, which forms atmosphere and setting.

<http://www.thepequod.org.uk/essays/litcrit/visualpo.htm>

- a) Write a paragraph in which you comment on content, style and language features that indicate that this is an academic text.
- b) Write a paragraph in the style of a film or literary critique stating which artistic genre (novels, theatre or film) is the best choice for getting a message across and why.

Oppgave 2 / Oppgave 2



<http://adland.tv/ooh/amnesty-international-ads-show-whats-happening> (adapted)

Write two or three paragraphs in which you interpret the message of the advertisement above, and discuss how the message is conveyed.

Oppgave 3 / Oppgave 3

a)

Write an interpretation of the poem "In Goya's Greatest Scenes We Seem to See..." by Lawrence Ferlinghetti ("Vedlegg 1") in which you also examine the poet's perspective on America. The pictures in "Vedlegg 2", two of which are paintings by Goya, may be used as inspiration for your answer.

b)

One of the literary genres you have studied is the short story.

Write a short story about the character in the lyrics "Misunderstood" ("Vedlegg 3") that reveals what type of person he is or represents. In your short story, you should make use of the features of this genre.

After the short story, write a separate paragraph explaining your reasons for the way your character thinks and acts in the short story, bearing in mind what the speaker reveals about himself in the lyrics.

c)

Read the extract from The Scarlet Letter in "Vedlegg 4".

Write a text in which you compare the extract's female character and her plight with another female character and her situation in a literary work or a film or play you have studied during your course. In your essay you should reflect on what these stories tell us about the situation women have lived in – or live in – and what these two characters teach us about gender roles.

d)

Writing about a book in the popular genre of vampire literature, a critic stated the following:

Just about all the interesting ideas that are introduced are immediately dropped and kicked out of sight. If there is anything at all that we can call a character, it is childlike and far too simplistic. In fact the characters are as vapid, as blank-eyed and expressionless as the author herself in one of her recent publicity photos. The plot hobbles along. All the same, the story could have been an admirable attempt at producing a "novel" in the tradition of *Dracula*, if only the author had not insisted on trying to emulate writers of better talent.

Another critic, Paul Jameson, strongly disagreed and remarked:

There is certainly literary merit in many of the novels in this genre, especially in the *Twilight* series.

Utdanningsdirektoratet

*Write an essay in which you **either** defend this literary genre, giving your reasons both for its literary value and its popularity, **or** write an essay in which you agree with those criticizing it and state what you consider to be good literature.*

Blank side

In Goya's Greatest Scenes We Seem to See ...

1 In Goya's greatest scenes we seem to see
the people of the world
exactly at the moment when
they first attained the title of
5 'suffering humanity'
They writhe upon the page
in a veritable rage
of adversity
Heaped up
10 groaning with babies and bayonets
under cement skies
in an abstract landscape of blasted trees
bent statues bats wings and beaks
slippery gibbets
15 cadavers and carnivorous cocks
and all the final hollering monsters
of the
'imagination of disaster'
they are so bloody real
20 it is as if they really still existed

And they do

Only the landscape is changed

They still are ranged along the roads
 plagued by legionnaires
 25 false windmills and demented roosters
 They are the same people
 only further from home
 on freeways fifty lanes wide
 on a concrete continent
 30 spaced with bland billboards
 illustrating imbecile illusions of happiness

The scene shows fewer tumbrils
but more strung-out citizens
in painted cars
and they have strange license plates
and engines
that devour America

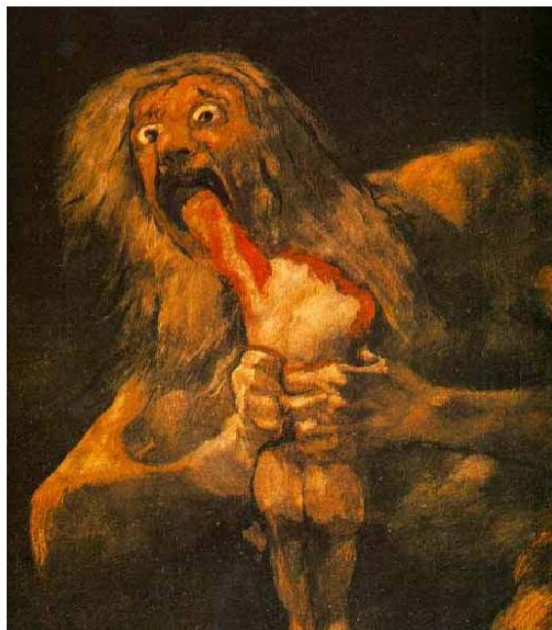
Lawrence Ferlinghetti

Ferlinghetti, L. 1958. *A Coney Island of the Mind, Poems by Lawrence Ferlinghetti*. New Directions Paperbook: pp.9-10.

Blank side



Bilde 1: Francisco Goya, "The Shootings of May 3, 1808" (1814) Oil on canvas, Museo del Prado, Madrid



Bilde 2: Francisco Goya, "Saturn Devouring One of His Sons" (1819-1823).
Oil mural transferred to canvas. Museo del Prado, Madrid.



Bilde 3: "Traffic"
<http://multivu.prnewswire.com/mnr/affinion/27862/images/27862-hi-traffic.jpg>

Blank side

Misunderstood

- 1 Just wanna be misunderstood
Wanna be feared in my neighbourhood
Just wanna be a moody man
Say things that nobody can understand
- 5 I wanna be obscure and oblique
Inscrutable and vague
So hard to pin down
I wanna leave open mouths when I speak
Want people to cry when I put them down
- 10 I wanna be either old or young
Don't like where I've ended up or where I begun
I always feel I must get things in the can
I just can't handle it the way I am
- 15 Why am I so straight and simple
People see through me like I'm made of glass
Why can't I deepen with greying temples
Am I growing out of my class
- 20 I always feel I should be somewhere else
I feel impatient like a girl on the shelf
They say that I should live sera sera¹⁾
But I am such an ordinary star
- 25 Coolwalkingsmoothtalkingstraightsmokingfirestoking
Coolwalkingsmoothtalking, yeah
- Just wanna be misunderstood
I wanna be feared in my neighbourhood
Just wanna be a moody man
- 30 Say things that nobody can understand
- Coolwalkingsmoothtalkingstraightsmokingfirestoking
Coolwalkingsmoothtalkingstraightsmokingfirestoking
Coolwalkingsmoothtalkingstraightsmokingfirestoking
Coolwalkingsmoothtalking, yeah

1) shortened version of *qué será, será*, which is Spanish for "what will be, will be".

Pete Townshend song lyrics, "Misunderstood", from the "Rough Mix" album, 1977.
http://www.lyricsmania.com/lyrics/pete_townshend_lyrics_4310/rough_mix_%5Bwith_ronnie_lane%5D_lyrics_13683/misunderstood_lyrics_158498.html

Blank side

The Scarlet Letter by Nathaniel Hawthorne

The story begins in Boston in the 17th century when Puritanism ruled. A young woman, Hester Prynne, gives birth to a daughter out of wedlock. She has to put the scarlet letter "A" on her bosom as punishment for adultery. Her husband sent her ahead to America but never arrived himself. It is believed that he has been lost at sea. Hester has apparently had an affair. She refuses to reveal her lover's identity, however, and the scarlet letter, along with public shaming, is the punishment for her alleged sin.

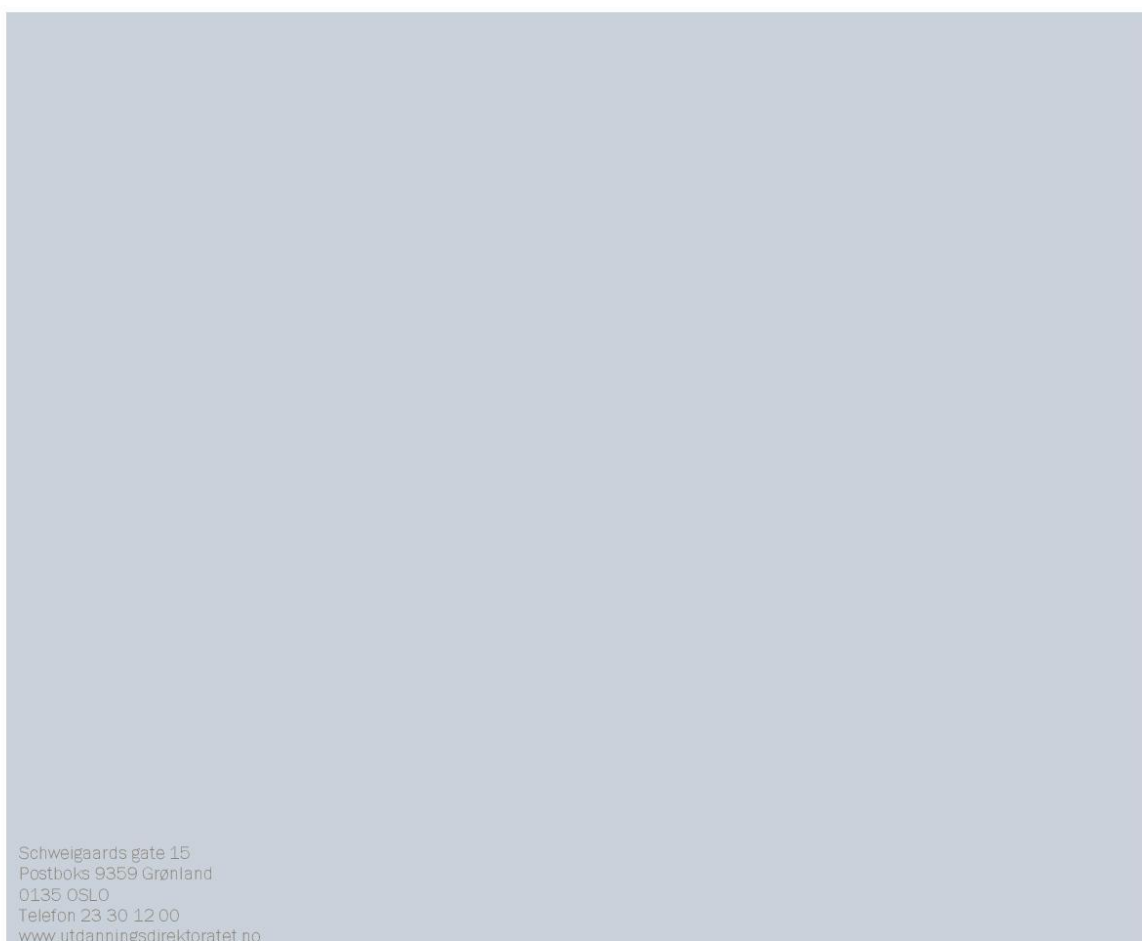
1 When the young woman – the mother of this child – stood fully revealed before the crowd, it seemed to be her first impulse to clasp the infant closely to her bosom; not so much by an impulse of motherly affection, as that she might thereby conceal a certain token, which was wrought or fastened into her dress. In a moment, however, wisely
5 judging that one token of her shame would but poorly serve to hide another, she took the baby on her arm, and, with a burning blush, and yet a haughty smile, and a glance that would not be abashed, looked around at her townspeople and neighbours. On the breast of her gown, in fine red cloth, surrounded with an elaborate embroidery and fantastic flourishes of gold thread, appeared the letter A. It was so artistically done, and with so
10 much fertility and gorgeous luxuriance of fancy, that it had all the effect of a last and fitting decoration to the apparel which she wore; and which was of a splendor in accordance with the taste of the age, but greatly beyond what was allowed by the sumptuary regulations of the colony.

The young woman was tall, with a figure of perfect elegance, on a large scale. She
15 had dark and abundant hair, so glossy that it threw off the sunshine with a gleam, and a face which, besides being beautiful from regularity of feature and richness of complexion, had the impressiveness belonging to a marked brow and deep black eyes. She was lady-like, too, after the manner of feminine gentility of those days; characterized by a certain state and dignity, rather than by the delicate, evanescent, and indescribable grace, which is now recognized as its indication. And never had Hester Prynne appeared more lady-like, in the antique interpretation of the term, than as she issued from the prison. Those who
20 had before known her, and had expected to behold her dimmed and obscured by a disastrous cloud, were astonished, and even startled, to perceive how her beauty shone out, and made a halo of the misfortune and ignominy in which she was enveloped. It may be true, that, to a sensitive observer, there was something exquisitely painful in it. Her attire, which, indeed, she had wrought for the occasion, in prison, and had modelled much after her own fancy, seemed to express the attitude of her spirit, the desperate recklessness of her mood, by its wild and picturesque peculiarity. But the point which
25 drew all eyes, and, as it were, transfigured the wearer, - so that both men and women, who had been familiarly acquainted with Hester Prynne, were now impressed as if they beheld her for the first time, - was that SCARLET LETTER, so fantastically embroidered and illuminated upon her bosom. It had the effect of a spell, taking her out of the ordinary relations with humanity, and inclosing her in a sphere by herself.

'She hath good skill at her needle, that's certain,' remarked one of the female
35 spectators; 'but did ever a woman, before this brazen hussy, contrive such a way of showing it! Why, gossips, what is it but to laugh in the faces of our godly magistrates, and make a pride out of what they, worthy gentlemen, meant for a punishment?'

'It were well,' muttered the most iron-visaged of the old dames, 'if we stripped
Madam Hester's rich gown off her dainty shoulders; and as for the red letter, which she
40 hath stitched so curiously, I'll bestow a rag of mine own rheumatic flannel, to make a fitter one!'

'O, peace, neighbours, peace!' whispered their youngest companion, 'Do not let her hear you! Not a stitch in that embroidered letter, but she has felt it in her heart.'



Schweigaards gate 15
Postboks 9359 Grønland
0135 OSLO
Telefon 23 30 12 00
www.undanningsdirektoratet.no



Appendix 6



Eksamen

31.05.2011

SPR3012 Engelskspråkleg litteratur og kultur
SPR3012 Engelskspråklig litteratur og kultur
Elevar/elever og privatistar/privatister

Nynorsk/Bokmål

Nynorsk

Eksamensinformasjon	
Eksamenstid	Eksamen varer i 5 timar.
Hjelpemiddel	Alle hjelpemiddel er tillatne, bortsett frå Internett og andre verktøy som kan brukast til kommunikasjon. Omsetjingsprogram er ikkje tillatne.
Bruk av kjelder	Dersom du bruker kjelder i svaret ditt, skal dei alltid førast opp på ein slik måte at lesaren kan finne fram til dei. Du skal føre opp forfattar og fullstendig tittel på både lærebøker og annan litteratur. Dersom du bruker utskrift eller sitat frå Internett, skal du føre opp nøyaktig nettadresse og nedlastingsdato.
Vedlegg	1. "Extract 1": <i>A Kestrel for a Knave</i> "Extract 2": <i>A Week in December</i> 2. "Far from me" "Let's Suppose" 3. <i>The Rug Merchant</i>
Informasjon om oppgåva	Du skal svare på alle tre oppgåvene. Oppgåve 1: Du skal svare på både deloppgåve a) og deloppgåve b). I deloppgåve a) skal du vise kompetansen din i å oppsummere synspunkta til ein forfattar i ein skjønnlitterær tekst. I deloppgåve b) skal du vise kompetansen din i å drøfte samanhengen mellom form, innhald og stilnivå i ein tekst. Oppgåve 2: Du skal svare på enten deloppgåve a) eller deloppgåve b). I deloppgåve a) skal du produsere ein tekst med tydelege sjangertrekk. I deloppgåve b) skal du tolke, analysere og vurdere eit dikt frå ein litteraturhistorisk periode ved å drøfte samanhengen mellom form, innhald og stilnivå i det. Oppgåve 3: Du skal svare på eitt av alternativa a), b), c) eller d). Oppgåve 3a prøver kompetansen din i å tolke og samanlikne synsvinkelen i to skjønnlitterære tekstar. Oppgåve 3b prøver kompetansen din i å tolke og samanlikne to litterære tekstar i eit kulturhistorisk og samfunnsmessig perspektiv. Oppgåve 3c prøver kompetansen din i å analysere to songtekstar ved å drøfte samanhengen mellom form, innhald og stilnivå i dei. Oppgåve 3d prøver kompetansen din i å analysere opningsdelen i ein skjønnlitterær tekst ved å drøfte kva for ein funksjon han har, kor effektiv han er og kva han fortel lesaren om personane i romanen og forholdet deira til kvarandre.

Kjelder	<p>Oppgåve 1: Shaw, G.B. 1111. "An Atrocious Institution". http://www.christmas-time.com/atrociousinstitution.htm. (13.02.2011).</p> <p>Oppgåve 2: Shaw, G.B. 1111. "An Atrocious Institution". http://www.christmas-time.com/atrociousinstitution.htm. (13.02.2011).</p> <p>Blake, W. 1794. "Infant Sorrow" frå diktsamlinga <i>Songs of Innocence and of Experience</i>. Oxford University Press: Oxford. Utgitt i nyutgåve av Triannon Press i 1967.</p> <p>Oppgåve 3a: Vedlegg 1: "Extract 1": Hines, B . 1969. <i>A Kestrel for a Knave</i>. Penguin Books: UK.</p> <p>Oppgåve 3b: Vedlegg 1: "Extract 1": Hines, B . 1969. <i>A Kestrel for a Knave</i>. Penguin Books: UK. "Extract 2": Faulks, S. 2009. <i>A Week in December</i>. Hutchinson: London.</p> <p>Oppgåve 3c: Vedlegg 2: "Far from me". http://staff.science.uva.nl/~jellekok/lyrics/boatman.html.</p> <p>"Let's Suppose". http://www.sing365.com/music/lyric.nsf/Love-At-First-Sight-lyrics-John-Mellencamp/AAB060E9E96095BB48257782000DF9A3.</p> <p>Oppgåve 3d: Vedlegg 3: Mullins, M. 2006. <i>The Rug Merchant</i>. Viking: New York.</p>
Informasjon om vurderinga	Sjå vurderingsrettleiinga med kjenneteikn på måloppnåing til sentralt gitt skriftleg eksamen. Vurderingsrettleiinga finn du på utdanningsdirektoratet.no.

Bokmål

Eksamensinformasjon	
Eksamenstid	Eksamen varer i 5 timer.
Hjelpemidler	<p>Alle hjelpemidler er tillatt, bortsett fra Internett og andre verktøy som kan brukes til kommunikasjon.</p> <p>Oversettelsesprogrammer er ikke tillatt.</p>
Bruk av kilder	<p>Hvis du bruker kilder i besvarelsen din, skal disse alltid oppgis på en slik måte at leseren kan finne fram til dem.</p> <p>Du skal oppgi forfatter og fullstendig tittel på både lærebøker og annen litteratur. Hvis du bruker utskrift eller sitat fra Internett, skal du oppgi nøyaktig nettside og nedlastingsdato.</p>
Vedlegg	<p>2. "Extract 1": <i>A Kestrel for a Knave</i> "Extract 2": <i>A Week in December</i></p> <p>2. "Far from me" "Let's Suppose"</p> <p>3. <i>The Rug Merchant</i></p>
Informasjon om oppgaven	<p>Du skal svare på alle tre oppgavene.</p> <p>Oppgave 1: Du skal besvare både deloppgave a) og deloppgave b). I deloppgave a) skal du vise din kompetanse i å oppsummere forfatterens synspunkter i en skjønnlitterær tekst. I deloppgave b) skal du vise din kompetanse i å drøfte sammenhengen mellom form, innhold og stilnivå i en tekst.</p> <p>Oppgave 2: Du skal besvare enten deloppgave a) eller deloppgave b). I deloppgave a) skal du produsere en tekst med tydelige sjangertrekk. I deloppgave b) skal du tolke, analysere og vurdere et dikt fra en litteraturhistorisk periode ved å drøfte sammenhengen mellom form, innhold og stilnivå i det.</p> <p>Oppgave 3: Du skal besvare ett av alternativene a), b), c) eller d).</p> <p>Oppgave 3a prøver din kompetanse i å tolke og sammenligne synsvinkelen i to skjønnlitterære tekster.</p> <p>Oppgave 3b prøver din kompetanse i å tolke og sammenligne to litterære tekster i et kulturhistorisk og samfunnsmessig perspektiv.</p> <p>Oppgave 3c prøver din kompetanse i å analysere to sangtekster ved å drøfte sammenhengen mellom form, innhold og stilnivå i dem.</p> <p>Oppgave 3d prøver din kompetanse i å analysere åpningsdelen i en skjønnlitterær tekst ved å drøfte hvilken funksjon den har, hvor effektiv den er og hva den forteller leseren om personene i romanen og deres forhold til hverandre.</p>

Kilder	<p>Oppgave 1: Shaw, G.B. 1111. "An Atrocious Institution". http://www.christmas-time.com/atrociousinstitution.htm. (13.02.2011).</p> <p>Oppgave 2: Shaw, G.B. 1111. "An Atrocious Institution". http://www.christmas-time.com/atrociousinstitution.htm. (13.02.2011).</p> <p>Blake, W. 1794. "Infant Sorrow" fra diktsamlingen <i>Songs of Innocence and of Experience</i>. Oxford University Press: Oxford. Utgitt i nyutgave av Triannon Press i 1967.</p> <p>Oppgave 3a: Vedlegg 1: "Extract 1": Hines, B. 1969. <i>A Kestrel for a Knave</i>. Penguin Books: UK.</p> <p>Oppgave 3b: Vedlegg 1: "Extract 1": Hines, B. 1969. <i>A Kestrel for a Knave</i>. Penguin Books: UK. "Extract 2": Faulks, S. 2009. <i>A Week in December</i>. Hutchinson: London.</p> <p>Oppgave 3c: Vedlegg 2: "Far from me". http://staff.science.uva.nl/~jellekok/lyrics/boatman.html.</p> <p>"Let's Suppose". http://www.sing365.com/music/lyric.nsf/Love-At-First-Sight-lyrics-John-Mellencamp/AAB060E9E96095BB48257782000DF9A3.</p> <p>Oppgave 3d: Vedlegg 3: Mullins, M. 2006. <i>The Rug Merchant</i>. Viking: New York.</p>
Informasjon om vurderingen	Se vurderingsveiledningen med kjennetegn på måloppnåelse til sentralt gitt skriftlig eksamen. Vurderingsveiledningen finner du på utdanningsdirektoratet.no .

Oppgave 1/Oppgave 1

The following text by George Bernard Shaw expresses what he thinks of Christmas.

Read the text and answer a) **and** b):

- a) Summarise in your own words the author's message in one or two sentences.
- b) Point out at least three different linguistic devices used by the author to argue his point, and explain how each of them enhances the message the author wants to convey.

An Atrocious Institution

Like all intelligent people, I greatly dislike Christmas. It revolts me to see every man rifle his neighbour's pockets under cover of a ghastly general pretence of festivity. It is really an atrocious institution, this Christmas. We must be gluttonous because it is Christmas. We must be drunken because it is Christmas. We must be insincerely generous; we must buy things that nobody wants, and give them to people we don't like; we must go to absurd entertainments, that make even our little children satirical; all because it is Christmas - that is, because the mass of the population, including the all-powerful middle-class tradesmen, depend on a week of licence and brigandage*, waste and intemperance to clear off its outstanding liabilities at the end of the year. As for me, I shall fly from it all tomorrow or next day to some remote spot miles from a shop, where nothing worse can befall me than a serenade from a few peasants. In town there is, for the moment, nothing for me or any honest man to do.

George Bernard Shaw

*brigandage *life and practice of brigands; highway robbery; plunder*

Oppgave 2/Oppgave 2

Answer either a) **or** b) below.

a) *Write a paragraph about another common festive event or holiday in which you attempt to convey an attitude similar to the one expressed by George Bernard Shaw in "An Atrocious Institution".*

b) A poem often develops one key idea, sentiment, emotion, picture, event or opinion.

Read the poem "Infant Sorrow" by William Blake and write two paragraphs in which you state what you think the speaker's view of the world is. In your answer you should reflect on how this is conveyed through linguistic devices and give examples of the language and style of the poem.

Infant Sorrow

My mother groaned, my father wept,
Into the dangerous world I leapt;
Helpless, naked, piping loud,
Like a fiend hid in a cloud.

Struggling in my father's hands,
Striving against my swaddling bands,
Bound and weary, I thought best
To sulk upon my mother's breast.

William Blake

Oppgave 3/Oppgave 3

Answer **one** of the questions 3a, 3b, 3c or 3d.

a)

Point of view is the perspective from which a story is told. In your studies you have learned about different points of view.

Write an essay in which you compare and contrast the point of view in "Extract 1": A Kestrel for a Knave ("Vedlegg 1") with the point of view in one of the English-language literary texts you have studied, discussing the effect created by the point of view in each text.

b)

Focus on the differences between the two schools depicted in "Extract 1": A Kestrel for a Knave (1968) and "Extract 2": A Week in December (2009) - both in "Vedlegg 1". Then write an essay in which you discuss what we can learn about developments in British society from the two texts and how you think British society will develop in the coming years.

c)

Love is the theme of many songs and is sometimes described in rosy clichés. However, love is often a more complicated matter.

Read the lyrics in "Vedlegg 2" and write an essay in which you analyse the lyrics of the two songs and discuss the views on love that each of them presents.

d)

"Vedlegg 3" presents the opening of Meg Mullins's novel *The Rug Merchant*.

Write an essay in which you discuss how effective the opening of the novel is, what we learn about the two characters Ushman Khan and Mrs Roberts and what we learn about the relationship between them. Use examples from the extract in your essay.

Extract 1: *A Kestrel for a Knave*

'Hymn number one-seven-five, "New every morning is the love".' The navy blue covers of the hymn books, inconspicuous against the dark shades of the boys' clothing, bloomed white across the hall as they were opened and the pages flicked through. The scuff and tick of the turning pages was slowly drowned under a rising chorus of coughing and hawking; until Mr Gryce, furious behind the lectern, scooped up his stick and began to smack it vertically down the face.

'STOP THAT INFERNAL COUGHING.'

The sight and swishsmack of the stick stopped the throat noises and the boys and the teachers, posted at regular intervals at the ends of the rows; all looked up to the platform. Gryce was straining over the top of the lectern like a bulldog up on its hind legs. 'It's every morning alike! As soon as the hymn is announced you're off revving up! Hm-hmm! Hm-hmm! It's more like a race track in here than an assembly hall! - hall - ringing across the hall, striking the windows and lingering there like the vibrations of a tuning fork.

No-one muffed. Not a foot scraped. Not a page stirred. The teachers looked seriously into the ranks of the boys. The boys stood looking up at Gryce, each one convinced that Gryce was looking at him.

The silence thickened. The boys began to swallow their adam's apples, their eyes skittering about in still heads. The teachers began to glance at each other and glance sideways up at the platform.

Then a boy coughed.

'Who did that?'

Everybody looking round.

'I said WHO DID THAT?'

The teachers moved in closer, alert like a riot squad.

'Mr. Crossley! Somewhere near you! Didn't you see the boy?'

Crossley flushed, and rushed amongst them, thrusting them aside in panic.

There Crossley! That's where it came from! Around there!' Crossley grabbed a boy by the arm and began to drag him into the open.

'It wasn't me, Sir!'

'Of course it was you.'

'It wasn't, Sir, honest!'

'Don't argue lad, I saw you.' Gryce thrust his jaw over the front of the lectern, the air whistling down his nostrils.

'MACDOWALL!' I might have known it! Get to my room, lad!'

(From *A Kestrel for a Knave* by Barry Hines, first published by Michael Joseph in 1968)

Extract 2: *A Week in December*

Radley Graves taught the lower years mostly, though he had a Year 11, GCSE, set as well. English had been fused with modern foreign languages and media studies under the banner of Communications, and this was something Radley felt he knew about. His training had focussed on the politics of race, gender and class, with hardly a mention of pupil management or lessons. These things had to be learned on the job, and initially Radley had found it difficult. At his first school, he'd twice been suspended and sent to anger management courses. Eventually, a senior colleague took pity and explained the principle of control. 'Never lock on. Never engage one-to-one. You can't win a battle of egos. Speak softly. Stay respectful. Answer without sarcasm, and if they won't co-operate, don't rebuke them. Never, ever raise your voice above a four out of ten.'

It took some practice. Radley didn't approve of the system because he felt it admitted that the teachers had ceded control to the pupils. The kids were allowed to come in late to lessons; and, after having once been threatened with a sexual harassment suit, Radley never again asked a girl pupil what had kept her. They were allowed to talk pretty much unchecked throughout a class, though if you could quell the noise by a soft and a generalised appeal, that was all right; what was not permitted was to single out the talker by name. They were allowed not to work if they didn't fancy it, though they could be gently reminded once. Swearing was permitted, unless it had racial or sexual overtones: Abir could call Radley himself a bastard but couldn't call Mehreen, who sat next to her, a bitch.

The corridors between lessons were a hard-hat zone, best avoided. The tiny first-year girls of barely four foot tall clung like shadows to the wall as raging fat six-footers of both sexes surged past them, heavy bags swinging, shouting down the length of the yellow and blue corridors. A bell rang, and Radley waited for the fireworks to stop going off in the stairwell before he made his way to the classroom. The moulded plastic chairs were set in twos at wipe-clean Formica-topped desks; only the carpet tiles, with brown mysterious stains, showed signs of the restless human traffic. [...]

(From *A Week in December* by Sebastian Faulks, 2009)

Far from Me

For you dear, I was born
For you I was raised up
For you I've lived and for you I will die
For you I am dying now
You were my mad little lover
In a world where everybody f***** everybody else over
You who are so far from me
Far from me
So far from me
Way across some cold neurotic sea
Far from me

I would talk to you of all matter of things
With a smile you would reply
Then the sun would leave your pretty face
And you'd retreat from the front of your eyes
I keep hearing that you're doing your best
I hope your heart beats happy in your infant breast
You are so far from me
Far from me
Far from me

There is no knowledge but I know it
There's nothing to learn from that vacant voice
That sails to me across the line
From the ridiculous to the sublime
It's good to hear you're doing so well
But really can't you find somebody else that you can ring and tell
Did you ever
Care for me?
Were you ever
There for me?
So far from me

You told me you'd stick by me
Through the thick and through the thin
Those were your very words
My fair-weather friend
You were my brave-hearted lover
At the first taste of trouble went running back to mother
So far from me
Far from me
Suspended in your bleak and fishless sea
Far from me
Far from me

Let's Suppose

Let's suppose you lived over here
Let's suppose I lived right over there
Let's suppose we shared a kiss
On a night like this
Let's suppose we went too far
In the back seat of your car
Let's suppose we fell in love
Like two turtle doves
And let's suppose we got engaged
Underneath that ol' Milky Way
And let's suppose we were complete
Wouldn't that be sweet
And let's suppose we found happiness
And built ourselves a nest
Let's suppose we had five kids
Like our grandparents did
And let's suppose our dreams came true
Just like they're supposed to do
And let's suppose we spent the rest of our lives together

But let's suppose sometimes we'd fight
When things weren't going right
Let's suppose you got tired of me
And our family
And let's suppose that we went broke
And we both just gave up hope
And let's suppose you left in a huff
Because you didn't love me enough
And let's suppose you found another man
And hit me in the head with a frying pan
And let's suppose I wound up dead
Underneath the bed
And what would happen to the kids
After all that we did
Let's suppose we both better just think it over

So I suppose it's just hello
And that it's time for us to go
We got a little ahead of ourselves
Like everybody else
Maybe some other time
I suppose that'd be just fine
So until we meet again
I suppose we're just friends
But remember what they say
'Cause it happens every day
I believe in love at first sight
So there, I said it
Good night

The Rug Merchant

Ushman Khan doesn't like tourists. It is June, though, and tourist season in New York, and, as he double-parks his van outside Mrs. Roberts's apartment building, they are everywhere. Dressed in T-shirts and sandals, fancy cameras dangling from their necks, they run for awnings and bus stops as hard rain begins to fall. Watching the chaos, Ushman smiles. He holds a newspaper over his head and, nearly dry, makes his way through the crowd. Mrs. Roberts's maid comes to the door; letting Ushman into the cool, dark apartment. The girl is young and squat and motions for Ushman to follow her into the den. He has been here before and walks silently through the expansive apartment. In nearly every room there is a rug from his small showroom on Madison Avenue. They are comforting to him, each one reminding him of his wife. She is still in Tabriz, where she selects and commissions hand-knotted wool and silk carpets and woven kilim from the women in the bazaar and private workshops. Every month she ships two or three rugs to Ushman with a note that reads, *Sell, Sell, Sell, Your Wife, Farak*.

Last year, when Mrs. Roberts and her husband moved across the park, she commissioned Ushman to cover all the floors in the new apartment. The job enabled Ushman to absorb the increase in the rent for his shop and nearly double his inventory, but he never liked being in Mrs. Roberts's apartment, with its high, echoing ceilings and drafty, perfumed air. And she had not been easy to please. Ushman must have hauled two dozen rugs into and out of the apartment after Mrs. Roberts had lived with them for a few days only to find their colors too muted, too bright, or just wrong. On these occasions, she seemed perversely pleased, as if she enjoyed finding flaws. And then, when Ushman rolled out the rug that was eventually to stay, Mrs. Roberts would appear in the doorway grimacing. Ushman, sweating and impatient, would ask, "Not this one, either?" She'd bite her bottom lip and place both hands on her hips. Then, as she turned away, she'd say, without looking back, "It's perfect. It stays." After he'd finished the project, Mrs. Roberts still kept in touch. Often, she would come by the shop with a friend of hers, winking at Ushman as the other woman knelt to admire an expensive silk. Or she would call, asking about a rug's origin or care or something else that Ushman knew was on the appraisal documents he'd given her. This persistence made him anxious, made him think that maybe she would try to say he had cheated her or been dishonest. So, when Mrs. Roberts called this morning and asked him to come to the apartment right away, Ushman felt nothing but dread. He waits for her in the den watching the rain fall in a dark haze over the park. It is an unusually cold day for late June and Ushman shivers, even in long sleeves. From a bedroom off the den, Mrs. Roberts appears just as Ushman is looking at his watch.

"You have somewhere to be," she says, more of a statement than a question.

"Yes, hello. An appointment." Then, with a gesture he learned from his father, Ushman raises his eyebrows and extends his arms toward her as if he were displaying a precious object. "But what may I do for you?"

Mrs. Roberts looks past him, out the window. "It's a terrible day to be working, Ushman," she says.

Ushman lowers his eyes, fiddles with the keys in his pocket. "You're right, a terrible day," he says.

Mrs. Roberts looks down at her feet. "I have grown a distaste for this rug," she says, walking across the center of it. "I want another."

Ushman smiles, tries hard to hide his irritation. "It's been over three months, Mrs. Roberts. The trial period is finished. You own the rug now."

"Oh, for goodness' sake, I know that. I don't want a refund, Ushman. This one will go to my niece, I think." She turns to look at the closed door from which she came, listens for a moment, and then turns back to Ushman. "But," she says, pacing the length of the rug. "I wanted to remind you of the space so you could pick out an appropriate few and show them to me."

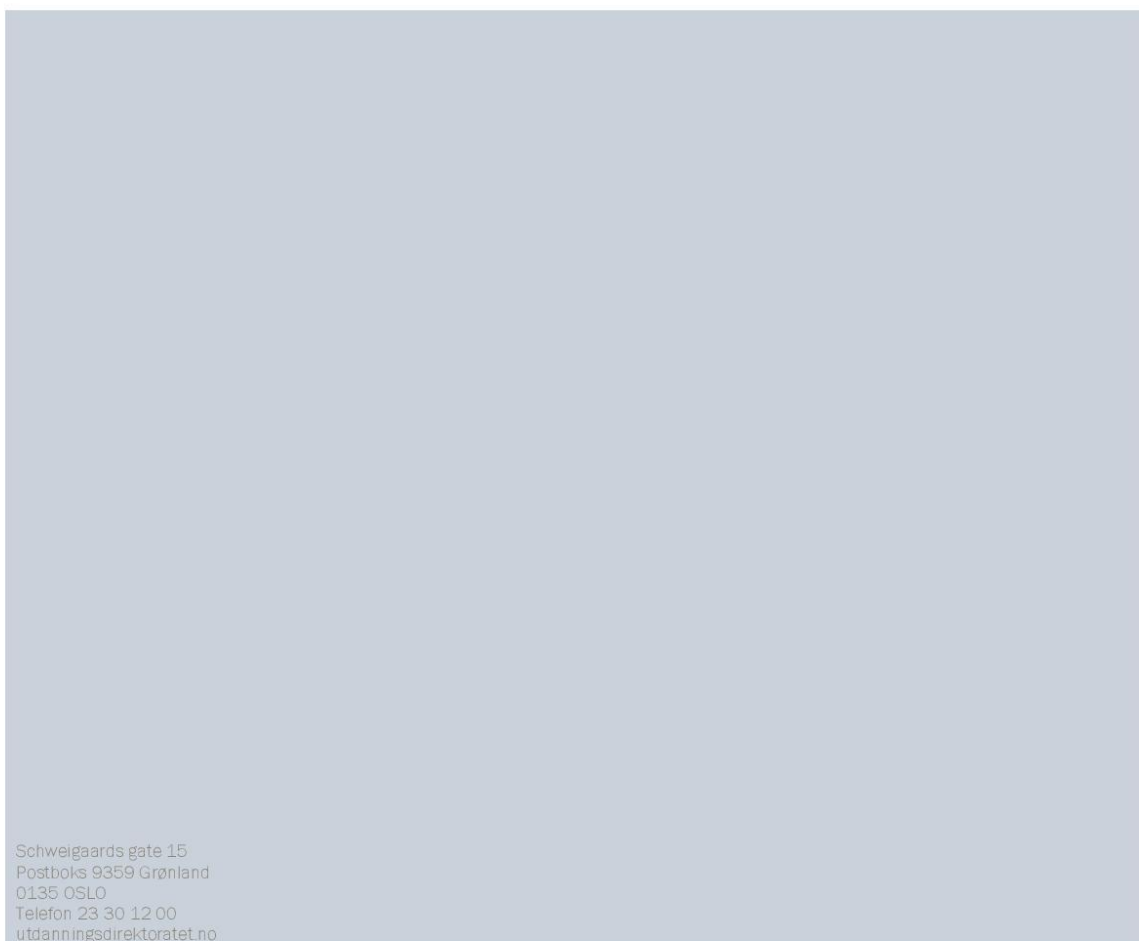
Ushman nods, but cannot contain his disapproval. The royal blue and red silk from Karaja truly belongs in the room. As any good rug should, it makes the room appear bigger, warmed, more textured. But, most importantly, it gives the room weight. "This rug, though, in this room, I have never seen a more perfect match...", he begins.

Mrs. Roberts waves her hand at Ushman, signaling for quiet.

(From *The Rug Merchant* by Meg Mullins, 2006)

Blank side

Blank side



Schweigaards gate 15
Postboks 9359 Grønland
0135 OSLO
Telefon 23 30 12 00
utdanningsdirektoratet.no

